

Coarse-graining Kohn-Sham Density Functional Theory

Phanish Suryanarayana^a, Kaushik Bhattacharya^{*,b}, Michael Ortiz^b

^aCollege of Engineering, Georgia Institute of Technology, GA 30332, USA

^bDivision of Engineering and Applied Science, California Institute of Technology, CA 91125, USA

Abstract

We present a real-space formulation for coarse-graining Kohn-Sham Density Functional Theory that significantly speeds up the analysis of material defects without appreciable loss of accuracy. The approximation scheme consists of two steps. First, we develop a linear-scaling method that enables the direct evaluation of the electron density without the need to evaluate individual orbitals. We achieve this by performing Gauss quadrature over the spectrum of the linearized Hamiltonian operator appearing in each iteration of the self-consistent field method. Building on the linear-scaling method, we introduce a spatial approximation scheme resulting in a coarse-grained Density Functional Theory. The spatial approximation is adapted so as to furnish fine resolution where necessary and to coarsen elsewhere. This coarse-graining step enables the analysis of defects at a fraction of the original computational cost, without any significant loss of accuracy. Furthermore, we show that the coarse-grained solutions are convergent with respect to the spatial approximation. We illustrate the scope, versatility, efficiency and accuracy of the scheme by means of selected examples.

Key words: Kohn-Sham; Density functional theory; Coarse-graining; Defects; Linear-scaling; Gauss quadrature

1. Introduction

Crystal defects play a critical role in determining macroscopic properties of solids even when present in small concentrations (Phillips, 2001). For instance, vacancies, perhaps the simplest type of defect, are fundamental to phenomena like creep, spall and radiation ageing even when present in only parts per million. Dislocations enable plasticity even though their densities are typically as small as 10^{-8} per atomic row. Stacking faults also influence plasticity and surface energies affect fracture. Further, small changes in composition often have a dramatic effect on mechanical properties because they segregate to defects. The reason that defects play such a profound role is that they bring together the chemical effects of the core, the discrete effects of the lattice and the long-range effects of the elastic fields. Consider specifically a vacancy. On the one hand, the nature of the core is determined by the chemistry of the dangling bonds. On the other hand, the vacancy gives rise to displacements of atoms that decay at a polynomial rate, and this enables long-range interactions with other defects as well as macroscopically applied fields. This coupling grows stronger in extended defects. Thus, a complete understanding of defects at physically relevant concentrations requires a simultaneous study of the details of the core as well as the long-range elastic fields.

This remains an outstanding challenge. On the one hand, methods like Density Functional Theory (DFT) that are capable of accurately describing the chemistry of the defect core are much too complicated

*Corresponding Author (bhatta@caltech.edu)

and expensive to use for defects with long-range fields. On the other hand, methods capable of describing the long-range fields including atomistic and continuum methods, lack the fidelity and rely on empiricism to describe the core. This challenge has prompted the development of multiscale approaches which seek to use DFT or tight-binding (TB) for the core, atomistic models for the near field and continuum for the far-field (e.g., Govind et al. (1999); Choly et al. (2005); Lu et al. (2006); Bernstein et al. (2009)). Others have used parameter passing where detailed simulations are used to fit the parameters of a coarser model (e.g., Cuitino et al. (2002)). While these multiscale methods provide valuable insight, there is no seamless transition from DFT to TB or empirical potentials to continuum. Uncontrolled approximations resulting from these seams, the use of linear-response theory and kinematic assumptions like the Cauchy-Born hypothesis may render the methods unreliable. Furthermore, there is often no guarantee of convergence of the solutions provided by these methods to the full DFT solution.

An alternative approach to multiscale modeling was initiated through the quasicontinuum method (Tadmor et al., 1996). Here, one seeks to solve a single theory in the complete region, but using an adaptive numerical method where all the details are represented close to the core but only sampled farther away. There is no additional modeling (uncontrolled approximations), and all the approximations are in the numerical scheme so that it converges to the detailed theory with increased resolution. Tadmor et al. (1996) and subsequent refinements (see Miller and Tadmor (2009) for a recent review) use an atomistic approach with empirical potentials as the only theory. To that extent, it is unable to describe the full details of the defect core.

The Schrödinger equation is fundamental for describing the quantum mechanical electronic structure of matter. However, the solution of the Schrödinger equation is exceedingly expensive (scaling exponentially with number of atoms) and therefore impractical. A far-reaching reformulation of the problem was achieved by Hohenberg and Kohn (1964), who proved the existence of a one-to-one correspondence between the ground-state electron density and the ground-state wavefunction of a many-particle system. By this correspondence, the electron density replaces the many-body electronic wavefunction as the fundamental unknown field, thereby greatly reducing the dimensionality and computational complexity of the problem. However, this theory requires an unknown but universal (independent of material) exchange and correlation functional. Computationally convenient models have been developed, including the local density approximation (LDA) (Kohn and Sham, 1965) and the generalized gradient approximation (GGA) (Langreth and Mehl, 1983; Perdew et al., 1992).

Traditional formulations of DFT follow Kohn and Sham (1965) and solve for the orbitals (Chelikowsky et al., 1994; Kresse and Furthmüller, 1996; Pask et al., 1999; Ismail-Beigi and Arias, 2000; Segall et al., 2002; Gonze et al., 2002; Tsuchida, 2004; Castro et al., 2006; Suryanarayana et al., 2010, 2011). Since the number of orbitals is proportional to the number of atoms and they have to be mutually orthogonal, these formulations typically result in cubic-scaling with respect to the number of atoms. This places severe restrictions on the size of the systems which can be studied. To overcome this, a number of methods have been proposed that exhibit better scaling properties, with particular emphasis on linear-scaling (Galli and Parrinello, 1992; Mauri et al., 1993; Goedecker, 1999; Skylaris et al., 2005; Barrault et al., 2007; Garcia-Cervera et al., 2009; Bowler and Miyazaki, 2012). See Goedecker (1999) and Bowler and Miyazaki (2012) for exhaustive reviews of these methods. Briefly, a vast majority of linear scaling methods avoid orbitals and instead reformulate DFT in terms of the so-called density matrix. This is a diagonally dominant matrix and linear scaling is obtained by cutting off the decaying off-diagonal components. Often a basis set made of localized Wannier functions is used to represent the density matrix, and the decay of off-diagonal components is equivalent to the decay of these functions. The cutoff is acceptable in insulators where the band gap results in a strong exponential decay. In metals however, one only has polynomial decay in general. While, it has been shown that exponential decay may exist in idealized examples at finite temperatures (Goedecker,

1999), a mathematical understanding of the decay properties and consequently the accuracy of these methods remain unclear. Therefore, the development of a linear-scaling method for metallic systems remains an open problem (Cances et al., 2008).

Even with a linear scaling DFT method, the study of crystal defects at realistic concentrations remains a daunting problem. Gavini et al. (2007) extended the quasicontinuum approach to electronic structure by implementing it on a simplified version of DFT known as orbital-free DFT (OFDFT). OFDFT models the kinetic energy of the electrons instead of computing it explicitly using wave functions, and is limited in its applicability to free electron metals like Aluminum. Gavini et al. (2007) use a nested discretization to fully resolve both the atomic displacements and electronic density close to the core, but sample it elsewhere. They demonstrated the efficacy of their method by exhibiting computational savings of a factor of 10^3 or more, and also showed through examples how the physics can change depending on the size of the computational cell (concentration of defects). Unfortunately, it is not possible to directly extend their idea to the standard formulation of DFT for three main reasons. First, orbitals are subject to orthogonality, which is a global constraint and difficult to coarse-grain. From a practical point of view, orthogonality means that orbitals oscillate extremely rapidly and this makes the quasicontinuum representation difficult. Second, they are not periodic even in homogenous systems (instead, they are Bloch-Floquet waves). Finally, one has as many orbitals as electrons, and thus the method would scale poorly even with a quasicontinuum representation of each orbital.

In this paper, we present a method that overcomes these difficulties. Specifically, we present a formulation (which we call CGDFT) that seamlessly coarse-grains DFT by recourse to controlled numerical approximations without the introduction of new or spurious physics. We accomplish this through a series of steps. First, we reformulate DFT in such a manner that eliminates the need to explicitly compute orbitals. Instead of computing the orbitals, and then using them to compute the quantities of interest including Fermi energy, electron density and band structure energy, we use integral representations of these quantities in terms of the projected density of states. Thus, the theory is formulated in terms of smoothly varying quantities that are amenable to coarse-graining. Second, we perform numerical Gauss quadrature on the spectrum of the Hamiltonian and use operator theory to evaluate these integrals. Third, we develop an algorithm that truly scales linearly with the number of atoms. Together we refer to these steps as the Linear Scaling Spectral Gauss Quadrature (LSSGQ) method. As a final step, the spatial approximation is adapted so as to furnish fine resolution where necessary and to coarsen elsewhere. This coarse-graining step enables the analysis of defects at a fraction of the original computational cost, without any significant loss of accuracy. Furthermore, we show that the CGDFT solutions are convergent with respect to the spatial approximation. These properties render CGDFT to be highly transferable, efficient and accurate.

The fundamental difference of the LSSGQ formulation from existing linear-scaling methods is that no explicit cutoff is employed. The formulation resembles the Fermi operator expansion (FOE) method (Goedecker, 1999; Bowler and Miyazaki, 2012), but there are some notable differences. Importantly, instead of expanding the density matrix in terms of a polynomial basis or rational functions, we use Gauss quadratures to evaluate diagonal elements of the density matrix and other quantities of interest. For this reason, LSSGQ does not involve the representation of the complete density matrix. Instead, only the components of interest are extracted. Further, the Fermi energy is not needed as an input to the method. This enables significant savings since the method does not have to be iterated over for the calculation of the Fermi energy. Additionally, symmetries in the system can be easily utilized to significantly reduce the computational effort. Finally the proposed method offers adaptivity through controllable accuracy in space, a useful feature for multiscale and coarse-graining formulations. The LSSGQ fomulation also has similarities to the recursion method used in tight-binding (Haydock et al., 1972, 1975; Haydock, 1980). In the recursion

method, a continued fraction representation of the projected density of states is developed and quantities of interest are obtained by integration over it. However, we proceed differently and directly calculate the Gauss quadrature rules, which is a much more accurate and stable process.

The remainder of this paper is organized as follows. In Section 2, we provide a brief introduction to DFT and in particular to the Kohn-Sham method. This includes a heuristic overview of our reformulation in Section 2.2. Next, we provide some mathematical background to spectral theory and Gauss quadrature in Section 3. Subsequently, we present and validate the LSSGQ formulation in Section 4. In Section 5, we describe CGDFT, and validate it through select examples. Finally, we conclude in Section 6.

2. Kohn-Sham Density Functional Theory

2.1. Orbital formulation

Consider a system of M atoms with N_e electrons. DFT is a theory which can be used to find the ground-state energy of the system, the ground-state electron density and the equilibrium position of the nuclei. In our presentation, we ignore electron spin and assume that N_e is even for clarity, but note that all the methods developed in the paper may easily be extended to include spin. Let $\mathbf{R} = \{\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_M\}$ denote the positions of the nuclei with charges $\{Z_1, Z_2, \dots, Z_M\}$ respectively. We follow [Kohn and Sham \(1965\)](#) and introduce orbitals $\Psi = \{\psi_1, \psi_2, \dots, \psi_{N_e/2}\}$ for the electrons. Then, the energy of the system is written as

$$\mathcal{E}(\Psi, \mathbf{R}) = - \sum_{n=1}^{N_e/2} \int_{\mathbb{R}^3} \psi_n^*(\mathbf{x}) \nabla^2 \psi_n(\mathbf{x}) d\mathbf{x} + E_{xc}(\rho) + E_H(\rho) + E_{\text{ext}}(\rho, \mathbf{R}) + E_{zz}(\mathbf{R}), \quad (1)$$

where the electron density

$$\rho(\mathbf{x}) = 2 \sum_{n=1}^{N_e/2} |\psi_n(\mathbf{x})|^2. \quad (2)$$

The first term in Eqn. 1 represents the kinetic energy of the non-interacting electrons. The second term, $E_{xc}(\rho)$, denotes the exchange-correlation energy. For definiteness, we adopt the local density approximation (LDA) ([Kohn and Sham, 1965](#))

$$E_{xc} = \int_{\mathbb{R}^3} \rho \epsilon_{xc}(\rho) d\mathbf{x} \quad (3)$$

where $\epsilon_{xc}(\rho)$ is the exchange-correlation energy density. The final three terms are electrostatic in nature,

$$E_H(\rho) = \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho(\mathbf{x}) \rho(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d\mathbf{x} d\mathbf{x}', \quad (4)$$

$$E_{\text{ext}}(\rho, \mathbf{R}) = \int_{\mathbb{R}^3} \rho(\mathbf{x}) \left(\sum_{I=1}^M \frac{Z_I}{|\mathbf{x} - \mathbf{R}_I|} \right) d\mathbf{x}, \quad (5)$$

$$E_{zz}(\mathbf{R}) = \frac{1}{2} \sum_{I=1}^M \sum_{\substack{J=1 \\ J \neq I}}^M \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}. \quad (6)$$

E_H is known as the Hartree energy and is the classical electrostatic interaction energy of the electron density, E_{ext} is the electrostatic interaction energy between the electron density and nuclear charges, and E_{zz} is the repulsive energy between the nuclei.

For a given position of the nuclei \mathbf{R} , the electronic ground-state energy of the system is obtained as the solution to the variational principle.

$$\mathcal{E}_0(\mathbf{R}) = \inf_{\Psi} \mathcal{E}(\Psi, \mathbf{R}) \quad (7)$$

subject to the orthonormality constraints

$$\int_{\mathbb{R}^3} \psi_i^*(\mathbf{x}) \psi_j(\mathbf{x}) d\mathbf{x} = \delta_{ij}, \quad i, j = 1, 2, \dots, N_e/2. \quad (8)$$

It is convenient to rewrite the electrostatic terms by introducing the electrostatic potential ϕ as the solution to the Poisson equation (cf e.g. [Ismail-Beigi and Arias \(2000\)](#); [Suryanarayana et al. \(2010\)](#))

$$-\frac{1}{4\pi} \nabla^2 \phi(\mathbf{x}, \mathbf{R}) = \rho(\mathbf{x}) + b(\mathbf{x}, \mathbf{R}) \quad (9)$$

where $b(\mathbf{x}, \mathbf{R}) = \sum_{J=1}^M b_J(\mathbf{x}, \mathbf{R}_J)$ denotes the total charge density of the nuclei, with $b_J(\mathbf{x}, \mathbf{R}_J)$ representing the regularized charge density of the J^{th} nucleus. Thereafter, the electrostatic energies may be rewritten as

$$\begin{aligned} E_H + E_{\text{ext}} + E_{\text{zz}} &= -\frac{1}{8\pi} \int_{\mathbb{R}^3} |\nabla \phi(\mathbf{x}, \mathbf{R})|^2 d\mathbf{x} + \int_{\mathbb{R}^3} (\rho(\mathbf{x}) + b(\mathbf{x}, \mathbf{R})) \phi(\mathbf{x}, \mathbf{R}) d\mathbf{x} \\ &\quad - \frac{1}{2} \sum_{J=1}^M \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{b_J(\mathbf{x}, \mathbf{R}_J) b_J(\mathbf{x}', \mathbf{R}_J)}{|\mathbf{x} - \mathbf{x}'|} d\mathbf{x} d\mathbf{x}'. \end{aligned} \quad (10)$$

where the last term denotes the self energy of the nuclei.

The Euler-Lagrange equations of the constrained variational principle (Eqns. 7 and 8) gives rise to the nonlinear eigenvalue problem

$$\mathcal{H} \psi_n = \lambda_n \psi_n, \quad n = 1, 2, \dots, N_e/2 \quad (11)$$

where the Hamiltonian

$$\mathcal{H} = -\frac{1}{2} \nabla^2 + V_{xc}(\rho) + \phi(\mathbf{x}, \mathbf{R}) \quad (12)$$

is a self-adjoint operator with eigenvalues λ_n (ordered so that $\lambda_1 \leq \lambda_2 \leq \dots$) and $V_{xc}(\rho) = \frac{\delta E_{xc}(\rho)}{\delta \rho(\mathbf{x})}$. The problem is nonlinear since the Hamiltonian depends on ρ (through $V_{xc}(\rho)$ and $\phi(\mathbf{x}, \mathbf{R})$) which in turn depends on ψ_n . Therefore, this problem is typically solved by a fixed point iteration with respect to the electron density, known as the self-consistent field (SCF) method (cf. e.g. [Martin \(2004\)](#)). In each iteration of the SCF method, the electron density is calculated by solving for the eigenfunctions ψ_n corresponding to the lowest $N_e/2$ eigenvalues λ_n , and then using Eqn. 2. This is indeed equivalent to the variational problem (cf. e.g. [Suryanarayana et al. \(2010\)](#)).

The electronic ground-state energy is obtained by substituting the solution of Eqn. 11 in Eqn. 1. We can then use the eigenvalue problem given by Eqn. 11 (left multiply by ψ_n^* and integrate) to show that this ground-state energy is

$$\begin{aligned} \mathcal{E}_0(\mathbf{R}) &= 2 \sum_{n=1}^{N_e/2} \lambda_n + E_{xc}(\rho) + \frac{1}{2} \int_{\mathbb{R}^3} (b(\mathbf{x}, \mathbf{R}) - \rho(\mathbf{x})) \phi(\mathbf{x}, \mathbf{R}) d\mathbf{x} \\ &\quad - \int_{\mathbb{R}^3} V_{xc}(\rho) \rho(\mathbf{x}) d\mathbf{x} - \frac{1}{2} \sum_{J=1}^M \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{b_J(\mathbf{x}, \mathbf{R}_J) b_J(\mathbf{x}', \mathbf{R}_J)}{|\mathbf{x} - \mathbf{x}'|} d\mathbf{x} d\mathbf{x}' \end{aligned} \quad (13)$$

where $U_{band} = 2 \sum_{n=1}^{N_e/2} \lambda_n$ denotes the band structure energy.

In this orbital approach, we solve the nonlinear eigenvalue problem (Eqn. 11) for the first $N_e/2$ eigenvalues and the corresponding eigenfunctions. Subsequently, the electronic ground-state energy can be evaluated using Eqn. 13. This formulation of DFT requires an effort scaling as $\mathcal{O}(N_e^3)$ (Goedecker (1999)), which restricts the size of systems that can be studied. Further, the orbitals are not amenable to coarse-graining since they are extremely oscillatory and are not periodic even in periodic systems. Finally, a quasicontinuum representation of each orbital would scale poorly since the number of orbitals increases with the number of electrons.

2.2. Heuristics of the proposed reformulation

The key idea behind our reformulation of DFT is to note that the ground-state energy (Eqn. 13) does *not* depend on individual eigenvalues and eigenfunctions, but only on certain sums:

$$U_{band} = 2 \sum_{n=1}^{N_e/2} \lambda_n, \quad \rho(\mathbf{x}) = 2 \sum_{n=1}^{N_e/2} |\psi_n(\mathbf{x})|^2. \quad (14)$$

So it seems natural to rewrite these sums by sampling and weighting:

$$U_{band} = 2 \sum_{k=1}^K w_k \bar{\lambda}_k, \quad \rho(\mathbf{x}) = 2 \sum_{k=1}^K w_k |\bar{\psi}_k(\mathbf{x})|^2 \quad (15)$$

for appropriately chosen sampling ‘eigenvalues’ $\bar{\lambda}_k$, ‘eigenfunctions’ $\bar{\psi}_k$ and weights w_k . We seek to find a way of choosing the sampling points optimally and computing the weights efficiently. Particularly, we would like K independent of N_e and $K \ll N_e/2$ (for large N_e) such that the overall effort in evaluating these sums is $\mathcal{O}(N_e)$. Since the number of electrons is proportional to the number of atoms, the computational effort will scale as $\mathcal{O}(M)$.

We accomplish this in four steps. We outline these here, and develop them precisely in Sections 4 after recalling some necessary mathematical preliminaries in Section 3. First, we introduce the orbital occupation function

$$g(\lambda, \lambda_f) = \begin{cases} 1, & \text{if } \lambda \leq \lambda_f \\ 0, & \text{otherwise} \end{cases} \quad (16)$$

and rewrite U_{band} and $\rho(\mathbf{x})$ as

$$U_{band} = 2 \sum_n \lambda_n g(\lambda_n, \lambda_f), \quad \rho(\mathbf{x}) = 2 \sum_n g(\lambda_n, \lambda_f) |\psi_n(\mathbf{x})|^2. \quad (17)$$

Above $\lambda_f = \lambda_{N_e/2}$ is called the Fermi energy. This is unknown *a priori* but can be solved through the constraint

$$N_e = 2 \sum_n g(\lambda_n, \lambda_f). \quad (18)$$

Second, we rewrite the sums in Eqns. 17 and 18 as spectral integrals of the form

$$I[f] = \int_{\sigma(\mathcal{H})} f(\lambda) d\sigma(\lambda) \quad (19)$$

where $\sigma(\mathcal{H})$ is the spectrum of the Hamiltonian \mathcal{H} and $d\sigma(\lambda)$ is the appropriate spectral measure. We provide an introduction to spectral theory in Section 3.1 and develop integral representations in Section 4.1. Third, we evaluate these spectral integrals using Gauss quadratures

$$I[f] \approx \sum_{k=1}^K w_k f(\lambda_k). \quad (20)$$

Finally, we evaluate the quadrature nodes $\{\lambda_k\}_{k=1}^K$ and weights $\{w_k\}_{k=1}^K$ using an efficient Lanczos type iteration. We provide an introduction to Gauss quadratures and the Lanczos type iteration in Section 3.2 and a precise algorithm in Section 4.2. We discuss the scaling and performance of the method in Section 4.3.

The reformulation as described above overcomes the difficulties encountered when attempting to coarse-grain the orbital formulation. Notably, the spectral quadrature points and weights can be formulated as local real space variables (at each spatial discretization point), and these are smoothly varying quantities. This enables us to coarse-grain the problem in Section 5.

2.3. Atomic positions

To calculate the ground-state energy of the system, we need to further minimize the energy with respect to the positions of the nuclei. To do so, we take the first variation of $\mathcal{E}_0(\mathbf{R})$ with respect to \mathbf{R} to obtain the forces on the nuclei

$$f_J = - \int_{\mathbb{R}^3} \frac{\partial b_J(\mathbf{x}, \mathbf{R}_J)}{\partial \mathbf{R}_J} (\phi(\mathbf{x}, \mathbf{R}) - \phi_J(\mathbf{x}, \mathbf{R}_J)) \, d\mathbf{x} \quad (21)$$

where $\phi_J(\mathbf{x}, \mathbf{R}_J) = \int_{\mathbb{R}^3} \frac{b_J(\mathbf{x}', \mathbf{R}_J)}{|\mathbf{x} - \mathbf{x}'|} \, d\mathbf{x}'$ and f_J represents the force on the J^{th} nucleus. In the special case where we have $b_J(\mathbf{x}, \mathbf{R}_J) = -Z_J \delta(\mathbf{x} - \mathbf{R}_J)$, we obtain

$$f_J = Z_J \nabla (\phi(\mathbf{x}, \mathbf{R}) - \phi_J(\mathbf{x}, \mathbf{R}_J)) \Big|_{\mathbf{x}=\mathbf{R}_J}. \quad (22)$$

This result is commonly referred to as the Hellmann-Feynman theorem (cf. e.g. [Finnis \(2003\)](#)).

2.4. Pseudopotential Approximation

The core states are localized in the vicinity of the nucleus leading to oscillations of the valence orbitals in this region. Irrespective of the basis set used, a large number of basis functions are required to capture these oscillations. Additionally, the tightly bound core electrons are chemically inactive and hence have a negligible contribution towards determining physical properties. In view of this, the core electrons are eliminated and an effective nuclear potential is introduced to describe the effect of the core electrons, resulting in nodeless pseudo-orbitals. This amounts to replacing the all-electron potential $V_{\text{ext}}(\mathbf{x}, \mathbf{R}) = \sum_{I=1}^M \frac{Z_I}{|\mathbf{x} - \mathbf{R}_I|}$ with an effective potential $V_{\text{ext}}^{\text{PS}}(\mathbf{x}, \mathbf{R})$.

The pseudopotentials can be broadly classified as either local or non-local based on their spatial dependence. The local pseudopotentials are explicit functions $V_{\text{ext}}^{\text{PS}}(\mathbf{x}, \mathbf{R})$. They can be incorporated into our formulation by letting $b(\mathbf{x}, \mathbf{R}) = \frac{-1}{4\pi} \nabla^2 V_{\text{ext}}^{\text{PS}}$. In contrast, non-local pseudopotentials are operators on the orbital with angular momentum dependence and are designed to accurately reproduce the scattering properties of the all-electron potential. These include norm conserving ([Bachelet et al., 1982](#); [Rappe et al., 1990](#); [Troullier and Martins, 1991](#)) and ultrasoft pseudopotentials ([Vanderbilt, 1990](#)), which are usually employed in the Kleinman-Bylander form ([Kleinman and Bylander, 1982](#)). For the examples in this paper, we utilize a local pseudopotential termed as the ‘Evanescent Core’ pseudopotential ([Fiolhais et al., 1995](#)). However, it should be noted that the methods proposed in this work are applicable even with the choice of non-local pseudopotentials.

2.5. Finite temperature extension

In previous discussion of the Kohn-Sham method, we have tacitly assumed a temperature of absolute zero. The method can be easily extended to finite temperatures (cf. e.g. [Parr and Yang \(1989\)](#)). The Kohn-Sham problem for the electronic ground-state takes the form

$$\begin{aligned}\mathcal{H}\psi_n &= \lambda_n\psi_n, \quad \mathcal{H} = -\frac{1}{2}\nabla^2 + V_{xc}(\rho) + \phi(\mathbf{x}, \mathbf{R}), \\ -\frac{1}{4\pi}\nabla^2\phi(\mathbf{x}, \mathbf{R}) &= \rho(\mathbf{x}) + b(\mathbf{x}, \mathbf{R}), \\ \rho(\mathbf{x}) &= 2\sum_n g(\lambda_n, \lambda_f) |\psi_n(\mathbf{x})|^2, \\ N_e &= 2\sum_n g(\lambda_n, \lambda_f)\end{aligned}\tag{23}$$

where the orbital occupation is now allowed to take fractional values as defined by the Fermi-Dirac distribution

$$g(\lambda, \lambda_f) = \frac{1}{1 + \exp(\frac{\lambda - \lambda_f}{\sigma})}.\tag{24}$$

Above, $\sigma = k_B\theta$, k_B being the Boltzmann constant and θ is the absolute temperature. The ground-state Helmholtz free energy

$$\mathcal{F} = \mathcal{E}_\sigma - \theta\mathcal{S}\tag{25}$$

where

$$\begin{aligned}\mathcal{E}_\sigma &= 2\sum_n g(\lambda_n, \lambda_f)\lambda_n + E_{xc}(\rho) + \frac{1}{2}\int_{\mathbb{R}^3} (b(\mathbf{x}, \mathbf{R}) - \rho(\mathbf{x}))\phi(\mathbf{x}, \mathbf{R}) \, d\mathbf{x} \\ &\quad - \int_{\mathbb{R}^3} V_{xc}(\rho)\rho(\mathbf{x}) \, d\mathbf{x} - \frac{1}{2}\sum_{J=1}^M \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{b_J(\mathbf{x}, \mathbf{R}_J)b_J(\mathbf{x}', \mathbf{R}_J)}{|\mathbf{x} - \mathbf{x}'|} \, d\mathbf{x}d\mathbf{x}'\end{aligned}$$

is the finite temperature counterpart of the ground-state energy at absolute zero (\mathcal{E}_0) and

$$\mathcal{S} = -2k_B \sum_n [g(\lambda_n, \lambda_f) \log g(\lambda_n, \lambda_f) + (1 - g(\lambda_n, \lambda_f)) \log(1 - g(\lambda_n, \lambda_f))]\tag{26}$$

is the entropy resulting from the fractional orbital occupations. The ground-state position of the nuclei can be found by equilibrating Eqn. 21 ([Weinert and Davenport, 1992](#)). Using the finite temperature calculation, it is also possible to get an accurate extrapolation for the ground-state energy at absolute zero ([Gillan, 1989](#))

$$\mathcal{E}_0 \approx \frac{1}{2} (\mathcal{E}_\sigma + \mathcal{F}).\tag{27}$$

3. Mathematical Background

In this section, we recall a few mathematical tools which we exploit to develop our methods. We refer the reader to [Rudin \(1991\)](#) for further details on spectral theory and [Golub and Meurant \(2010\)](#) for Gauss quadratures.

3.1. Spectral theory

Let \mathcal{H} be a self-adjoint linear operator on a finite-dimensional Hilbert space H with inner product (\cdot, \cdot) and norm $\|\cdot\|$. Then, according to the spectral theorem, we have a unique resolution of the identity \mathcal{E} that satisfies

$$\mathcal{H} = \int_{\sigma(\mathcal{H})} \lambda d\mathcal{E}(\lambda) \quad (28)$$

where $\sigma(\mathcal{H}) = \{\lambda_1, \lambda_2, \dots, \lambda_{N_d}\} \subset \mathbb{R}$ is the spectrum of \mathcal{H} . Consequently, for any bounded function f on $\sigma(\mathcal{H})$,

$$f(\mathcal{H}) = \int_{\sigma(\mathcal{H})} f(\lambda) d\mathcal{E}(\lambda). \quad (29)$$

Note that Eqn. 29 can also be written using Dirac's bra-ket notation as $f(\mathcal{H}) = \sum_{n=1}^{N_d} f(\lambda_n) |\psi_n\rangle \langle \psi_n|$, where ψ_n denotes the eigenfunction corresponding to the eigenvalue λ_n . We assume that the eigenvalues are ordered such that $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_{N_d}$.

Let $\{\eta_p\}_{p=1}^{N_d}$ be an orthonormal basis of H so that for any $\zeta \in H$, $\zeta = \sum_{p=1}^{N_d} \zeta_p \eta_p$. The eigenfunctions can therefore be expressed as $\psi_n = \sum_{p=1}^{N_d} \psi_{n,p} \eta_p$. Thereafter, we have the following representation of the measure $\mathcal{E}_{\zeta, \zeta}(\lambda)$ on $\sigma(\mathcal{H})$

$$\mathcal{E}_{\zeta, \zeta}(\lambda) = (\mathcal{E}(\lambda)\zeta, \zeta) = \begin{cases} 0, & \text{if } \lambda < \lambda_1 \\ \sum_{n=1}^m \sum_{p=1}^{N_d} \sum_{q=1}^{N_d} \psi_{n,p} \psi_{n,q} \zeta_p \zeta_q, & \text{if } \lambda_m \leq \lambda < \lambda_{m+1} \\ \sum_{n=1}^{N_d} \sum_{p=1}^{N_d} \sum_{q=1}^{N_d} \psi_{n,p} \psi_{n,q} \zeta_p \zeta_q, & \text{if } \lambda_{N_d} < \lambda \end{cases} \quad (30)$$

In the special case of $\zeta = \eta_k$, it reduces to

$$\mathcal{E}_{\eta_k, \eta_k}(\lambda) = \begin{cases} 0, & \text{if } \lambda < \lambda_1 \\ \sum_{n=1}^m |\psi_{n,k}|^2, & \text{if } \lambda_m \leq \lambda < \lambda_{m+1} \\ \sum_{n=1}^{N_d} |\psi_{n,k}|^2, & \text{if } \lambda_{N_d} < \lambda. \end{cases} \quad (31)$$

Therefore for any $\zeta \in H$,

$$(f(\mathcal{H})\zeta, \zeta) = \int_{\sigma(\mathcal{H})} f(\lambda) d\mathcal{E}_{\zeta, \zeta}(\lambda) = \int_a^b f(\lambda) d\mathcal{E}_{\zeta, \zeta}(\lambda), \quad (32)$$

for $a = \lambda_1$ and $b = \lambda_{N_d}$. It is worth noting that $d\mathcal{E}_{\zeta, \zeta}(\lambda)$ represents the projected density of states of \mathcal{H} .

3.2. Gauss quadrature

Let \mathcal{P} be the space of real polynomials and \mathcal{P}_K be a subspace of \mathcal{P} consisting of polynomials of degree K . We define an inner product (relative to the measure $\mathcal{E}_{\zeta, \zeta}$) of two polynomials $p, q \in \mathcal{P}$ as

$$\langle p, q \rangle_{\zeta} = \int_a^b p(\lambda) q(\lambda) d\mathcal{E}_{\zeta, \zeta}(\lambda) \quad (33)$$

with norm

$$\|p\|_{\zeta} = \left(\int_a^b p^2(\lambda) d\mathcal{E}_{\zeta, \zeta}(\lambda) \right)^{\frac{1}{2}}. \quad (34)$$

Here, we are interested in approximating integrals of the form (Eqn. 32)

$$I[f] = \int_a^b f(\lambda) d\mathcal{E}_{\zeta, \zeta}(\lambda) \quad (35)$$

using a quadrature rule. To do so, we approximate the function $f(\lambda)$ with an interpolation polynomial

$$f(\lambda) \approx \sum_{k=1}^K f(\lambda_k^\zeta) l_k^\zeta(\lambda) \quad (36)$$

where $\{\lambda_k^\zeta\}_{k=1}^K$ are the nodes/interpolation points and $l_k^\zeta(\lambda)$ is the Lagrange polynomial

$$l_k^\zeta(\lambda) = \prod_{j=1, j \neq k}^K \frac{\lambda - \lambda_j^\zeta}{\lambda_k^\zeta - \lambda_j^\zeta}. \quad (37)$$

Thereafter, we obtain the quadrature formula

$$I[f] \approx \sum_{k=1}^K w_k^\zeta f(\lambda_k^\zeta) \quad (38)$$

where

$$w_k^\zeta = \int_a^b l_k^\zeta(\lambda) d\mathcal{E}_{\zeta, \zeta}(\lambda) \quad (39)$$

are the weights of the quadrature rule. In order to obtain the quadrature rule of highest degree, we further consider the nodes $\{\lambda_k\}_{k=1}^K$ as unknowns. This results in a quadrature rule of degree $2K - 1$ i.e. any $p \in \mathcal{P}_{2K-1}$ is integrated exactly, commonly referred to as Gauss quadrature.

To efficiently evaluate the nodes and weights of the Gauss quadrature rule, we generate a sequence of orthonormal polynomials (with respect to the measure $\mathcal{E}_{\zeta, \zeta}$) $\{\hat{p}_k\}_{k=0}^K$ through the three-term recurrence relationship

$$\begin{aligned} b_{k+1} \hat{p}_{k+1}(\lambda) &= (\lambda - a_{k+1}) \hat{p}_k(\lambda) - b_k \hat{p}_{k-1}(\lambda), \quad k = 0, 1, \dots, K-1 \\ \hat{p}_{-1}(\lambda) &= 0, \quad \hat{p}_0(\lambda) = 1, \quad b_0 = 1 \end{aligned} \quad (40)$$

where

$$a_{k+1} = \langle \lambda \hat{p}_k, \hat{p}_k \rangle_\zeta, \quad k = 0, 1, \dots, K-1 \quad (41)$$

and \hat{b}_k is computed such that $\|\hat{p}_k\|_\zeta = 1$, $k = 0, 1, \dots, K$. Corresponding to these orthonormal polynomials, there is a tridiagonal Jacobi matrix \hat{J}_K of dimension K

$$\hat{J}_K = \begin{pmatrix} a_1 & b_1 & & & \\ b_1 & a_2 & b_2 & & \\ & \ddots & \ddots & \ddots & \\ & & b_{K-2} & a_{K-1} & b_{K-1} \\ & & & b_{K-1} & a_K \end{pmatrix} \quad (42)$$

Let us denote $\hat{P}_K(\lambda) = (\hat{p}_0(\lambda), \hat{p}_1(\lambda), \dots, \hat{p}_{K-1}(\lambda))^T$. Then the three term recurrence relation given by Eqn. 40 can be written compactly as

$$\lambda \hat{P}_K(\lambda) = \hat{J}_K \hat{P}_K(\lambda) + b_K \hat{p}_K(\lambda) e_K \quad (43)$$

where e_K is the last column of the identity matrix of dimension K . It follows that the eigenvalues of \hat{J}_K (which are also the zeros of $\hat{p}_K(\lambda)$) are the nodes $\{\lambda_k\}_{k=1}^K$ of the Gauss quadrature rule and the weights $\{w_k\}_{k=1}^K$ are the squares of the first elements of the normalized eigenvectors.

4. Linear-Scaling Spectral Gauss Quadrature Method

In this section, we describe the formulation, implementation and validation of the proposed linear-scaling spectral Gauss quadrature (LSSGQ) method. The outline of this section is as follows. We first develop integral representations for the quantities of interest in Section 4.1. Subsequently, we describe the LSSGQ method in Section 4.2 and discuss its scaling and performance in Section 4.3. Next, we discuss the LSSGQ method in the context of the finite-difference approximation in Section 4.4. Finally, we validate the LSSGQ method through examples in Section 4.5. The connections of the LSSGQ method with the recursion method (Haydock (1980)) and the Padé approximation are discussed in Appendix A.

4.1. Integral representations

Consider a spatial discretization of our domain with N_d basis functions. Let \mathcal{H} be the (linearized) Hamiltonian (Eqn. 12) with a fixed ρ and ϕ . In order to solve the Kohn-Sham problem, we need to evaluate the Fermi energy, updated electron density, band structure energy and entropy. We begin by providing an integral representation of these quantities over the spectrum of \mathcal{H} .

4.1.1. Fermi energy

The Fermi energy is calculated by solving for the constraint

$$N_e = 2 \sum_{n=1}^{N_d} g(\lambda_n, \lambda_f). \quad (44)$$

Since $\|\psi_n\| = \sum_{p=1}^{N_d} |\psi_{n,p}|^2 = 1$, it follows that

$$\begin{aligned} \sum_{n=1}^{N_d} g(\lambda_n, \lambda_f) &= \sum_{p=1}^{N_d} \sum_{n=1}^{N_d} g(\lambda_n, \lambda_f) |\psi_{n,p}|^2 \\ &= \sum_{p=1}^{N_d} \int_a^b g(\lambda, \lambda_f) d\mathcal{E}_{\eta_p, \eta_p}(\lambda). \end{aligned} \quad (45)$$

Therefore, Eqn. 44 can be rewritten as

$$N_e = 2 \sum_{p=1}^{N_d} \int_a^b g(\lambda, \lambda_f) d\mathcal{E}_{\eta_p, \eta_p}(\lambda). \quad (46)$$

4.1.2. Electron density

The electron density at any point \mathbf{x}_0 is

$$\begin{aligned}
\rho(\mathbf{x}_0) &= 2 \sum_{n=1}^{N_d} g(\lambda_n, \lambda_f) |\psi_n(\mathbf{x}_0)|^2 \\
&= 2 \sum_{n=1}^{N_d} \sum_{p=1}^{N_d} \sum_{q=1}^{N_d} g(\lambda_n, \lambda_f) \psi_{n,p} \psi_{n,q} \eta_p(\mathbf{x}_0) \eta_q(\mathbf{x}_0) \\
&= 2 \int_a^b g(\lambda, \lambda_f) d\mathcal{E}_{\zeta, \zeta}(\lambda).
\end{aligned} \tag{47}$$

where $\zeta = \sum_{p=1}^{N_d} \eta_p(\mathbf{x}_0) \eta_p$.

4.1.3. Band structure energy

The band structure energy may be expressed as

$$\begin{aligned}
U_{band} &= 2 \sum_{n=1}^{N_d} \lambda_n g(\lambda_n, \lambda_f) \\
&= 2 \sum_{p=1}^{N_d} \sum_{n=1}^{N_d} \lambda_n g(\lambda_n, \lambda_f) |\psi_{n,p}|^2 \\
&= 2 \sum_{p=1}^{N_d} \int_a^b \lambda g(\lambda, \lambda_f) d\mathcal{E}_{\eta_p, \eta_p}.
\end{aligned} \tag{48}$$

4.1.4. Entropy

Finally, the entropy follows as

$$\begin{aligned}
S &= -2k_B \sum_{n=1}^{N_d} [g(\lambda_n, \lambda_f) \log g(\lambda_n, \lambda_f) + (1 - g(\lambda_n, \lambda_f)) \log(1 - g(\lambda_n, \lambda_f))] \\
&= -2k_B \sum_{p=1}^{N_d} \sum_{n=1}^{N_d} [g(\lambda_n, \lambda_f) \log g(\lambda_n, \lambda_f) + (1 - g(\lambda_n, \lambda_f)) \log(1 - g(\lambda_n, \lambda_f))] |\psi_{n,p}|^2 \\
&= -2k_B \sum_{p=1}^{N_d} \int_a^b [g(\lambda, \lambda_f) \log g(\lambda, \lambda_f) + (1 - g(\lambda, \lambda_f)) \log(1 - g(\lambda, \lambda_f))] d\mathcal{E}_{\eta_p, \eta_p}.
\end{aligned} \tag{49}$$

4.2. Formulation

We proceed to describe the proposed LSSGQ method. We solve the nonlinear Kohn-Sham eigenvalue problem using the SCF method (cf., e. g., [Martin \(2004\)](#)). Customarily, in each iteration of the SCF method, the eigenvalues and eigenfunctions of the Hamiltonian (\mathcal{H}) are evaluated for a given electron density. The updated electron density is evaluated using Eqn. 2, and the process is repeated until convergence. However, we circumvent the evaluation of the eigenfunctions and directly evaluate the electron density by employing the procedure described below.

We seek to evaluate the integral representations of the Fermi energy and electron density described earlier using Gauss quadratures. However, only the operator \mathcal{H} is known and its resolution of the identity $\mathcal{E}(\lambda)$ is unknown *a priori*. In order to overcome this difficulty, we use the spectral theorem to rewrite the recurrence relation given by Eqn. (40) as the Lanczos-type iteration

$$\begin{aligned} b_{k+1}v_{k+1} &= (\mathcal{H} - a_{k+1})v_k - b_kv_{k-1}, \quad k = 0, 1, \dots, K-1 \\ v_{-1} &= 0, \quad v_0 = \zeta, \quad b_0 = 1 \end{aligned} \quad (50)$$

where

$$a_{k+1} = (\mathcal{H}v_k, v_k), \quad k = 0, 1, \dots, K-1 \quad (51)$$

and b_k is computed such that $\|v_k\| = 1$, $k = 0, 1, \dots, K-1$. Note that the initial condition $v_0 = \zeta$ is chosen depending on the measure $\mathcal{E}_{\zeta, \zeta}(\lambda)$ with respect to which integration needs to be performed. Thereafter, the nodes and corresponding weights of the quadrature rule are ascertained following the procedure described in Section 3.2. Therefore, we arrive at

$$N_e = 2 \sum_{p=1}^{N_d} \int_a^b g(\lambda, \lambda_f) d\mathcal{E}_{\eta_p, \eta_p}(\lambda) \approx 2 \sum_{p=1}^{N_d} \sum_{k=1}^K w_k^{\eta_p} g(\lambda_k^{\eta_p}, \lambda_f) \quad (52)$$

$$\rho(\mathbf{x}_0) = 2 \int_a^b g(\lambda, \lambda_f) d\mathcal{E}_{\zeta, \zeta}(\lambda) \approx 2 \sum_{k=1}^K w_k^{\zeta} g(\lambda_k^{\zeta}, \lambda_f) \quad (53)$$

where $\zeta = \sum_{p=1}^{N_d} \eta_p(\mathbf{x}_0)\eta_p$.

Once the self-consistent solution of the Kohn-Sham problem for a given position of the nuclei is calculated, the free energy (Eqn. 25) is calculated by using Gauss quadrature for the evaluation of the band structure energy and entropy

$$U_{band} = 2 \sum_{p=1}^{N_d} \int_a^b \lambda g(\lambda, \lambda_f) d\mathcal{E}_{\eta_p, \eta_p}(\lambda) \approx 2 \sum_{p=1}^{N_d} \sum_{k=1}^K w_k^{\eta_p} \lambda_k^{\eta_p} g(\lambda_k^{\eta_p}, \lambda_f), \quad (54)$$

$$\begin{aligned} S &= -2k_B \sum_{p=1}^{N_d} \int_a^b [g(\lambda, \lambda_f) \log g(\lambda, \lambda_f) + (1 - g(\lambda, \lambda_f)) \log(1 - g(\lambda, \lambda_f))] d\mathcal{E}_{\eta_p, \eta_p}(\lambda) \\ &\approx -2k_B \sum_{p=1}^{N_d} \sum_{k=1}^K w_k^{\eta_p} [g(\lambda_k^{\eta_p}, \lambda_f) \log g(\lambda_k^{\eta_p}, \lambda_f) + (1 - g(\lambda_k^{\eta_p}, \lambda_f)) \log(1 - g(\lambda_k^{\eta_p}, \lambda_f))]. \end{aligned} \quad (55)$$

To calculate the ground-state free energy, we need to further minimize the free energy with respect to the positions of the nuclei. To this end, we equilibrate the forces on the nuclei given by Eqn. 21. We summarize the LSSGQ method in Algorithm 1.

4.3. Scaling and Performance

In the LSSGQ method, the Fermi energy, electron density, band structure energy and entropy are evaluated by performing Gauss quadrature on their respective integral representations. The cost of evaluating each integral is determined by the calculation of the matrix \hat{J}_K (Eqn. 42) via the recurrence relation (Eqn. 50)

Algorithm 1: LSSGQ method

Generate guess for positions of the nuclei (\mathbf{R})

repeat Relaxation of atoms

 Calculate charge density of the nuclei (b)

 Generate guess for electron density (ρ)

repeat Self-Consistent loop: SCF

 Calculate electrostatic potential (ϕ) by solving Eqn. 9

 Calculate exchange correlation potential (V_{xc})

 Calculate spectral Gauss quadrature nodes and weights

 Calculate Fermi energy (λ_f) using Eqn. 52

 Calculate electron density (ρ) using Eqn. 53

 Update the electron density (mixing)

until *Convergence of self-consistent iteration;*

 Calculate the forces on the nuclei using Eqn. 21

until *Energy minimized with respect to positions of atoms;*

Evaluate free energy \mathcal{F} using Eqns. 54 and 55 for the band structure energy and entropy respectively.

and the evaluation of its eigenvalues and eigenvectors. The size of \hat{J}_K in turn is ascertained by the number of nodes used for the quadrature (K). We claim that the number of quadrature nodes required for achieving a predefined accuracy is independent of system size based on the following argument in the spirit of Goedecker (1999). For definiteness, consider the evaluation of the electron density where the function to be integrated is $g(\lambda, \lambda_f)$. Since there are K quadrature nodes, the resolution offered is roughly proportional to $1/K$. Therefore, we can roughly estimate

$$K \propto \frac{\lambda_{N_d} - \lambda_1}{\Delta\lambda}, \quad (56)$$

where $\Delta\lambda$ is the width of the region where $g(\lambda, \lambda_f)$ takes fractional values. The ratio in Eqn. 56 is independent of system size, since the number of basis functions required per atom is usually independent of system size. Further, if the orthonormal basis used to discretize H is localized, then the matrix representation of \mathcal{H} is sparse. In such a situation, the evaluation of the quadrature nodes and weights is independent of system size. Since we have to evaluate $\mathcal{O}(N_d)$ such quadrature rules, and since N_d scales as number of atoms M , we conclude that the entire method scales as $\mathcal{O}(M)$. This is indeed verified by the results presented in Section 4.5.

To understand the performance of the LSSGQ method, consider any integral of the form given by Eqn. 35. The error incurred by using Gauss quadrature is given by

$$R[f] = \|p_K\|_{\zeta}^2 \frac{f^{2K}(c)}{(2K)!} \quad (57)$$

for some $c \in [a, b]$. It is to be expected that the accuracy of the LSSGQ method improves with increasing the temperature θ . Further, the LSSGQ method is also expected to have spectral convergence with respect to the number of quadrature points (K). These properties are verified through the test cases presented in Section 4.5.

In this work, we have developed quadrature rules of degree $2K - 1$ by treating both the nodes and weights as unknowns, thus performing Gauss quadrature. It is also possible to fix the positions of the nodes

and treat only the weights as unknowns. However, this will result in a quadrature rule of degree $K - 1$. Therefore, our choice of Gauss quadrature is optimal in the sense that for a given number of nodes, it has the highest degree quadrature rule. Also, in the proposed approach, the Fermi energy is not needed as an input, which highlights the efficiency of the method.

4.4. Finite-difference approximation

The method developed above is general and can be implemented with any localized orthogonal basis. We now briefly describe the implementation with the finite-difference approximation. This provides a concrete example of the method, and also prepares for the examples that follow. Consider a domain in dimension D , and discretize it with a uniform finite-difference grid with nodes $\{\mathbf{x}^p\}_{p=1}^{N_d}$ and spacing h . By discretizing the Laplacian in the Hamiltonian using finite-differences of chosen order, we obtain the matrix eigenvalue problem

$$\mathbf{H}\psi_n = \lambda_n\psi_n, \quad n = 1, 2, \dots, N_d \quad (58)$$

where $\psi_n = [\psi_{n,1}, \psi_{n,2}, \dots, \psi_{n,N_d}]^T$ is the n^{th} eigenvector or orbital and $\psi_{n,p}$ is the value of the orbital at node p . The Lanczos iteration (Eqn. 50) takes the form

$$\begin{aligned} b_{k+1}\mathbf{v}_{k+1} &= (\mathbf{H} - a_{k+1})\mathbf{v}_k - b_k\mathbf{v}_{k-1}, \quad k = 0, 1, \dots, K-1 \\ \mathbf{v}_{-1} &= \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{v}_0 = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad b_0 = 1 \end{aligned} \quad (59)$$

where

$$a_{k+1} = \mathbf{v}_k^T \mathbf{H} \mathbf{v}_k, \quad k = 0, 1, \dots, K-1 \quad (60)$$

and b_k is computed such that $\|\mathbf{v}_k\|_{l^2} = 1$. For the p^{th} finite-difference node, only the p^{th} component of the initial vector \mathbf{v}_0 is nonzero and is set to unity. The calculation of the spectral quadrature nodes $\{\lambda_k^p\}_{k=1}^K$ and weights $\{w_k^p\}_{k=1}^K$ for each node $p \in [1, N_d]$ proceeds as described in Section 4.2. Importantly, note that these values may be computed independently at each node. Once the spectral quadrature nodes and weights are known, we evaluate the Fermi energy by solving for the constraint

$$N_e = h^D \sum_{p=1}^{N_d} \rho^p \quad (61)$$

where

$$\rho^p = \frac{2}{h^D} \sum_{k=1}^K w_k^p g(\lambda_k^p, \lambda_f) \quad (62)$$

is the electron density at the p^{th} finite-difference node. Furthermore, the band structure energy and entropy can be evaluated

$$U_{band} = \sum_{p=1}^{N_d} u^p, \quad (63)$$

$$S = \sum_{p=1}^{N_d} s^p \quad (64)$$

where

$$u^p = 2 \sum_{k=1}^K w_k^p \lambda_k^p g(\lambda_k^p, \lambda_f), \quad (65)$$

$$s^p = -2k_B \sum_{k=1}^K w_k^p [g(\lambda_k^p, \lambda_f) \log g(\lambda_k^p, \lambda_f) + (1 - g(\lambda_k^p, \lambda_f)) \log(1 - g(\lambda_k^p, \lambda_f))]. \quad (66)$$

The above equations allow us to interpret u^p and s^p as the nodal band structure energy and nodal entropy.

4.5. Numerical Examples and Validation

In this section, we proceed to verify the LSSGQ method through selected examples. Our first example concerns a one-dimensional model problem proposed by [García-Cervera et al. \(2009\)](#). Since the problem shares many of the features of the linearized Kohn-Sham problem, it provides a convenient test of the accuracy and performance of the LSSGQ method. In our second example, we apply the LSSGQ method to the nonlinear three-dimensional Kohn-Sham problem and evaluate selected crystal properties. Finally, we showcase the scope and versatility of the LSSGQ method by studying the phenomenon of surface relaxation.

4.5.1. One-dimensional model problem

Consider a one-dimensional chain of atoms positioned with unit spacing at $\{R_J\}_{J=1}^M$. Let the effective potential due to an atom at R_J be given by ([García-Cervera et al., 2009](#))

$$V_J(x) = -\frac{\alpha}{\sqrt{2\pi\beta^2}} \exp\left(-\frac{(x - R_J)^2}{2\beta^2}\right) \quad (67)$$

and therefore the total potential at any point is given by $V(x) = \sum_{J=1}^M V_J(x)$. Here, α and β represent the depth and width of the wells respectively. The ground-state properties are determined through the eigenvalue problem

$$\mathcal{H}\psi_n = \lambda_n \psi_n, \quad n = 1, 2, \dots \quad (68)$$

where the Hamiltonian is given by

$$\mathcal{H} = -\frac{1}{2} \frac{d^2}{dx^2} + V(x). \quad (69)$$

The Fermi energy, electron density and energy are calculated using the relations

$$N_e = \sum_{n=1}^{N_d} g(\lambda_n, \lambda_f), \quad (70)$$

$$\rho(x) = \sum_{n=1}^{N_d} g(\lambda_n, \lambda_f) |\psi_n(x)|^2, \quad (71)$$

$$\mathcal{E}_\sigma = \sum_{n=1}^{N_d} g(\lambda_n, \lambda_f) \lambda_n. \quad (72)$$

We introduce a 12th order accurate finite-difference approximation for the Laplacian and use the LSSGQ method described in Section 4.4. By appropriately choosing the parameters α and β , we can vary the band gap of the system, thereby choosing between insulating, semiconducting and metallic systems. Further

details can be found in [García-Cervera et al. \(2009\)](#). We consider two sets of parameters $(\alpha, \beta) = (100, 0.3)$ and $(\alpha, \beta) = (10, 0.45)$, which we designate as insulator and metal, respectively. We also study the effect of temperature on the efficacy of the method, by considering two extreme cases corresponding to $\sigma = 0.0001$ and $\sigma = 1.0$. In the results presented below, the error in energy is defined as the difference between the energies obtained by the LSSGQ method and diagonalization. Similarly, error in electron density is the L^2 norm of the difference in electron densities obtained on using the LSSGQ method and diagonalization. We note that all errors have been normalized by the corresponding quantities obtained via diagonalization.

Non-periodic calculations. For these calculations, we choose a chain of 101 atoms, i.e. $M = 101$. The convergence of the LSSGQ method with respect to the number of spectral quadrature points for both metals and insulators is presented in Fig. 1. The rapid convergence of the energy for both metals and insulators is clear from the figure. For insulators, the convergence is independent of temperature, a consequence of the presence of a band gap. By contrast, convergence for metals is significantly accelerated on increasing the temperature.

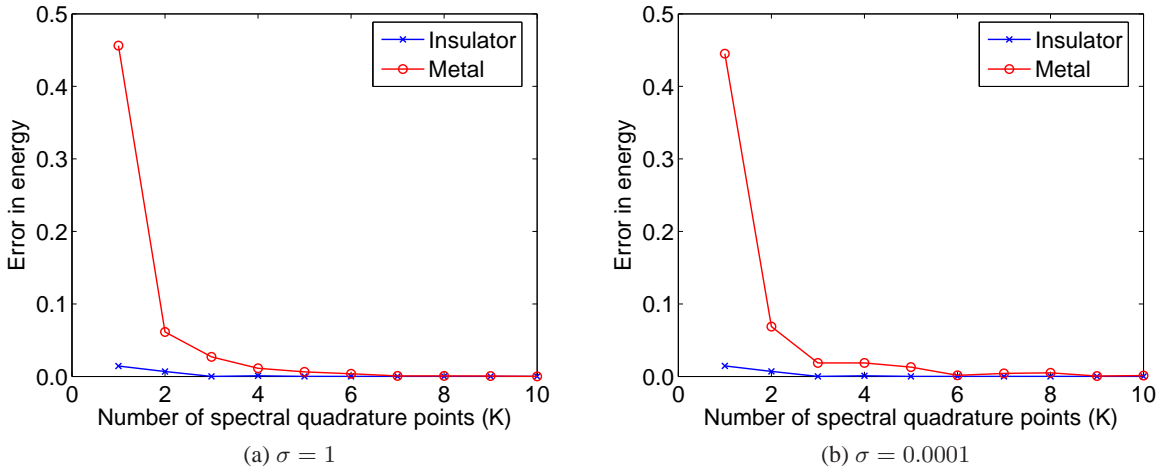


Figure 1: Convergence of the LSSGQ method for the one-dimensional model problem

Next, we introduce a defect by removing the center atom and evaluate the defect energy as well as the defect electron density. The defect energy and defect electron density are defined as the difference in energies and electron densities between systems with and without the defect. Accurately predicting these quantities is challenging, as it requires the calculation of relatively small differences. The results so obtained are presented in Fig. 2. For the insulator, there is rapid convergence with number of spectral quadrature points and the convergence is independent of the temperature. However, for the metal we notice that the number of spectral quadrature points required is dependent on the temperature, with fewer points required on increasing the temperature.

Periodic calculations. Consider next an infinite chain of atoms with unit spacing. We define a two atom unit cell as the representative volume Ω_{RV} . Since the potential decays exponentially, we evaluate the potential inside Ω_{RV} by considering the contribution of atoms which are within a cutoff radius. We then follow the procedure outlined in Appendix B to evaluate the Fermi energy, electron density and energy per atom. The results so obtained are presented in Fig. 3. As observed in the non-periodic calculations, convergence in

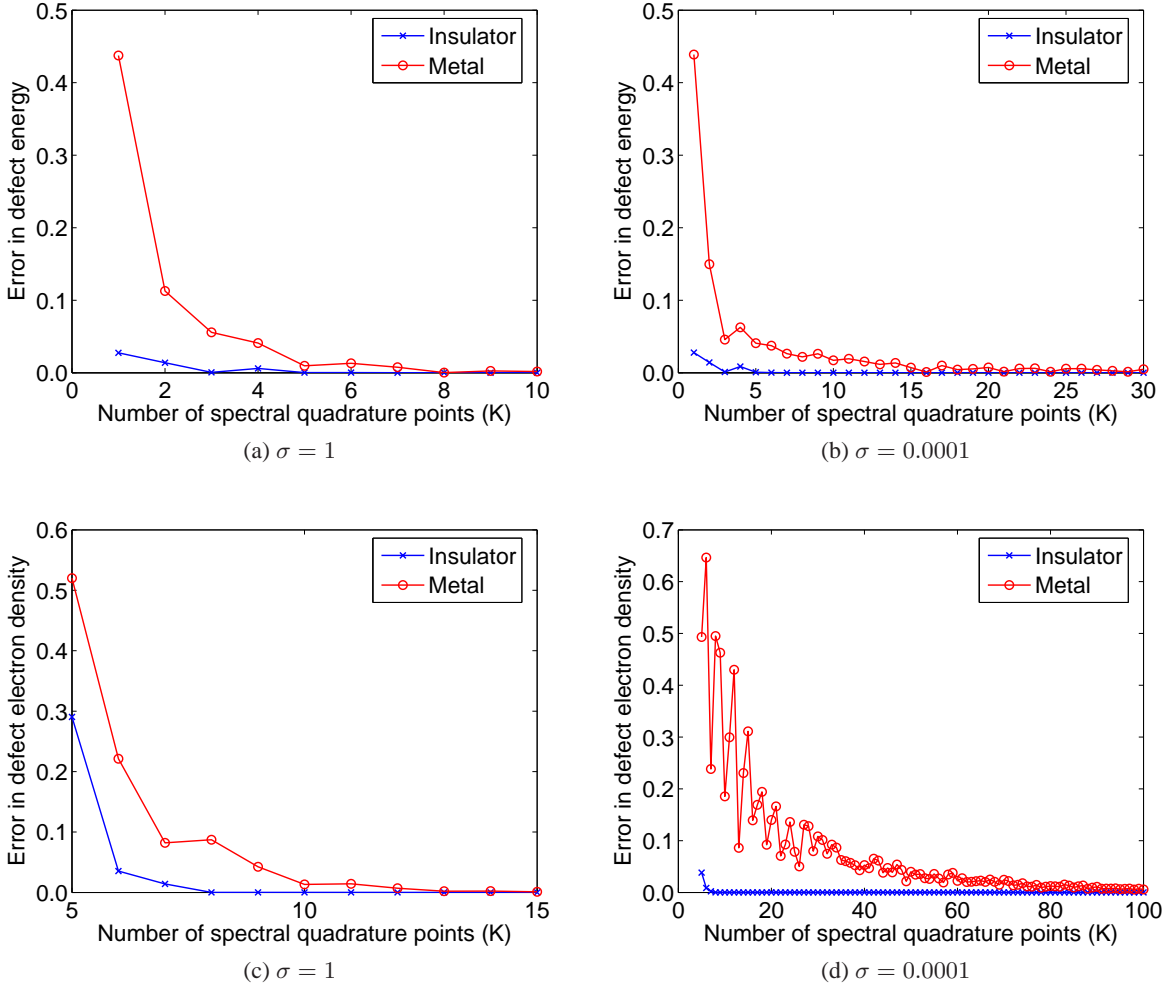


Figure 2: Convergence of the defect energy and defect electron density for the one-dimensional model problem

insulators is independent of temperature, whereas for metals convergence is significantly accelerated with increasing temperature.

Performance and Scaling. We now examine the rate of convergence as well as the scaling properties of the LSSGQ method. First, we perform non-periodic calculations on a chain of 101 atoms ($M = 101$). We plot the convergence in energy with respect to the number of spectral quadrature points in Fig. 4. It is clear that we get spectral convergence, as is to be expected, since Gauss quadrature is itself a spectrally convergent method.

Next, we study the scaling of the LSSGQ method in terms of the computational time taken versus the system size. We perform non-periodic calculations for metallic systems ranging from 1 atom to 100,000 atoms, all on a single computer. This has been showcased in Fig. 5, which highlights that the LSSGQ method is indeed linear-scaling. It should be noted that such large systems are intractable for orbital formulations, especially when solving on a single computer.

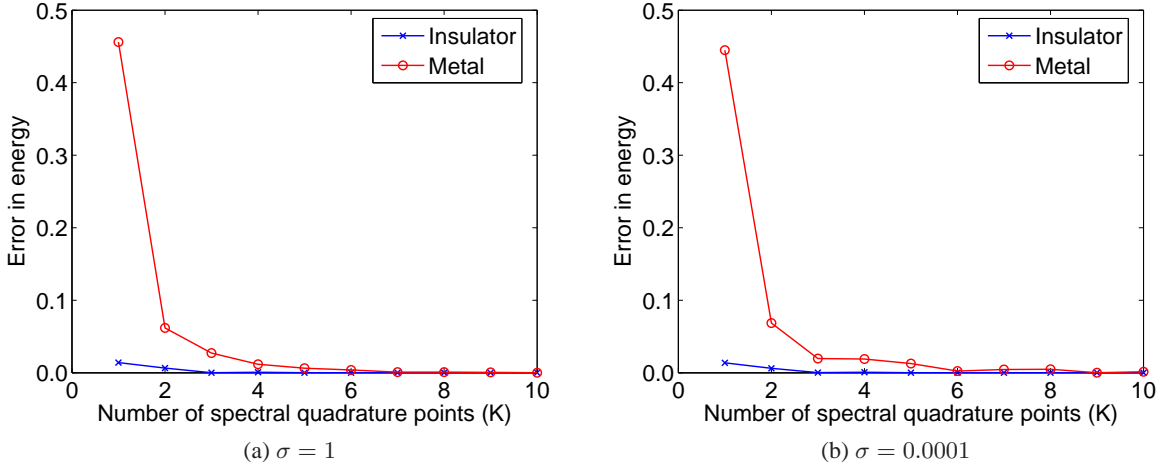


Figure 3: Convergence of the LSSGQ method for the periodic one-dimensional model problem

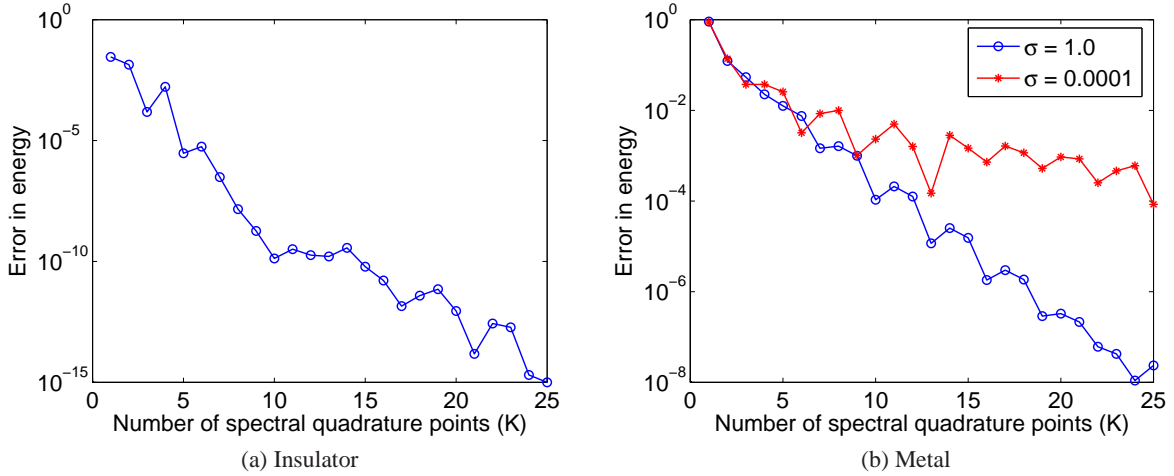


Figure 4: Spectral convergence of the LSSGQ method

4.5.2. Kohn-Sham problem

In the previous section, we have verified the LSSGQ method by means of a one-dimensional chain of atoms interacting through a model potential. In our next example, we apply the LSSGQ method to the solution of the Kohn-Sham problem using the LDA (Kohn and Sham (1965)) for the exchange correlation energy, and the ‘Evanescent Core’ pseudopotential approximation (Fiolhais et al., 1995). Specifically, we evaluate the bulk properties of body centered cubic (BCC) crystals of sodium, lithium and study the phenomenon of (001) surface relaxation of BCC sodium.

In our calculations, we utilize the procedure outlined in Algorithm 1 to solve for the ground-state properties. Specifically, we use a 6th order accurate finite-difference discretization of the Laplacian. We solve for the equilibrium atomic positions through the BFGS quasi-Newton method with a cubic line search procedure (cf., e. g., Vogel (2002)). For the SCF method, we use the generalized Broyden method (Fang and Saad,

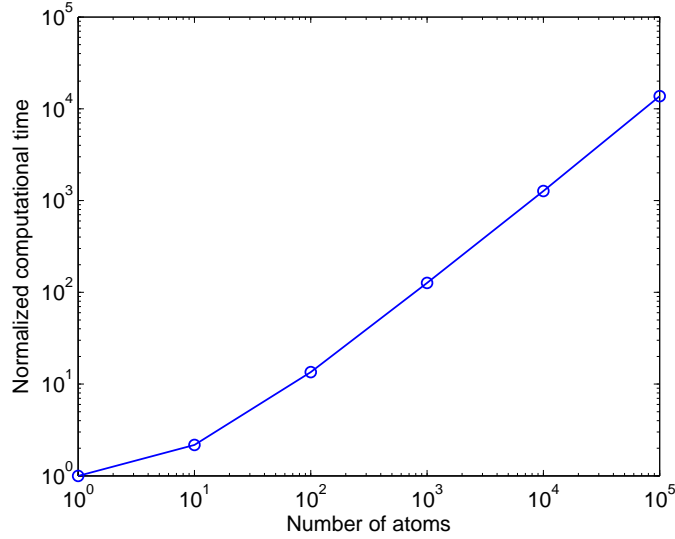


Figure 5: Linear-scaling nature of the LSSGQ method

2009) for convergence acceleration. We handle the periodicity in the system for the eigenvalue problem through the procedure outlined in Appendix B. We calculate the Fermi energy by using a combination of bisection, secant and inverse quadratic interpolation methods (Forsythe et al., 1973) to solve Eqn. 52. We solve the Poisson equation (Eqn. 9) with the generalized minimal residual method (GMRES - Saad and Schultz (1986)), wherein we evaluate the charge density $b(\mathbf{x}, \mathbf{R})$ by employing a prespecified radius. In order to evaluate the free energy of the system, we need to perform spatial integrations. To this end, we employ the trapezoidal rule, wherein each quantity is assumed to be constant in a cube of side h around each finite-difference point. Here, h denotes the spacing of the finite-difference nodes. Finally, we choose $\sigma = 0.8$ eV for all our calculations and extrapolate to find the ground-state energy at absolute zero using Eqn. 27.

Crystal properties. We begin by evaluating the bulk properties of BCC sodium and lithium. We designate the BCC unit cell as the representative volume and use periodic boundary conditions. First, we verify the convergence of the LSSGQ method. To do so, we plot the binding energy per atom as a function of the number of spectral quadrature points (K) in Fig. 6. It is clear that we obtain rapid convergence, highlighting the efficacy of the LSSGQ method. Next, we plot the binding energy per atom as a function of lattice constant in Fig. 7. Using a cubic fit to this data, we calculate the cohesive energy, equilibrium lattice constant and bulk modulus. The results so obtained are presented in Tables 1 and 2 with a comparison to previous results obtained by Fiolhais et al. (1995). There is good agreement in the cohesive energy and the lattice constant. However, there is a reasonable disagreement in the bulk modulus. The fact that we obtain a smooth curve for the binding energy versus the lattice constant (Fig. 7) gives us confidence in our results. Nevertheless, the discrepancy warrants further investigation.

Surface relaxation. Next we study the (001) surface relaxation of BCC sodium. We choose the representative volume (Ω_{RV}) to be a rectangular cuboid of dimensions $a \times a \times h$, where ‘ a ’ is the lattice constant. Furthermore, $h = h_{vac} + h_{cell}$, where $h_{vac} = n_1 a$ and $h_{cell} = n_2 a$ ($n_1, n_2 \in \mathbb{Z}$) are the heights of vacuum and material included in Ω_{RV} , respectively. We use periodic boundary conditions in the x_1, x_2 directions. In addition, we use zero Dirichlet boundary conditions on the top face of Ω_{RV} and assume that the perfectly

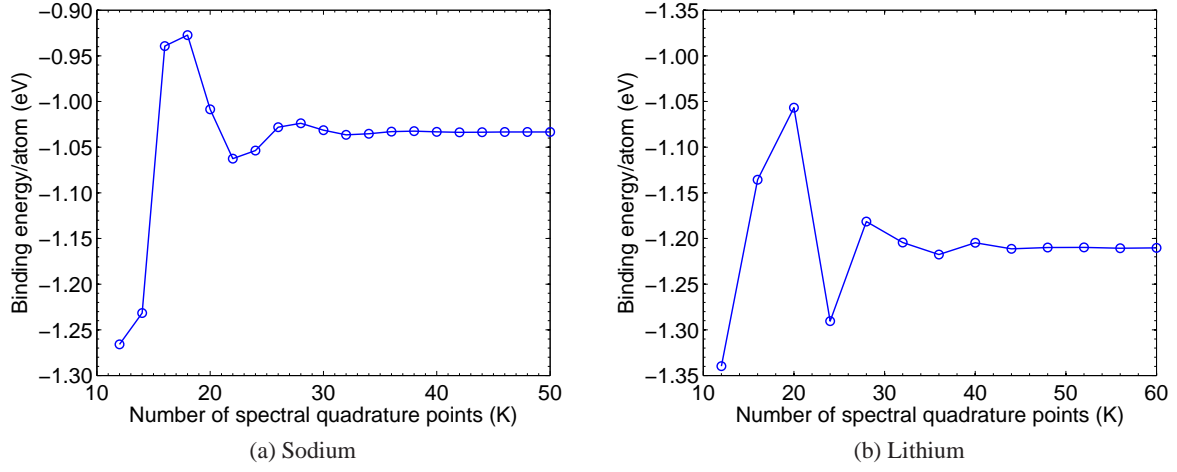


Figure 6: Convergence of the LSSGQ method for evaluating the crystal properties

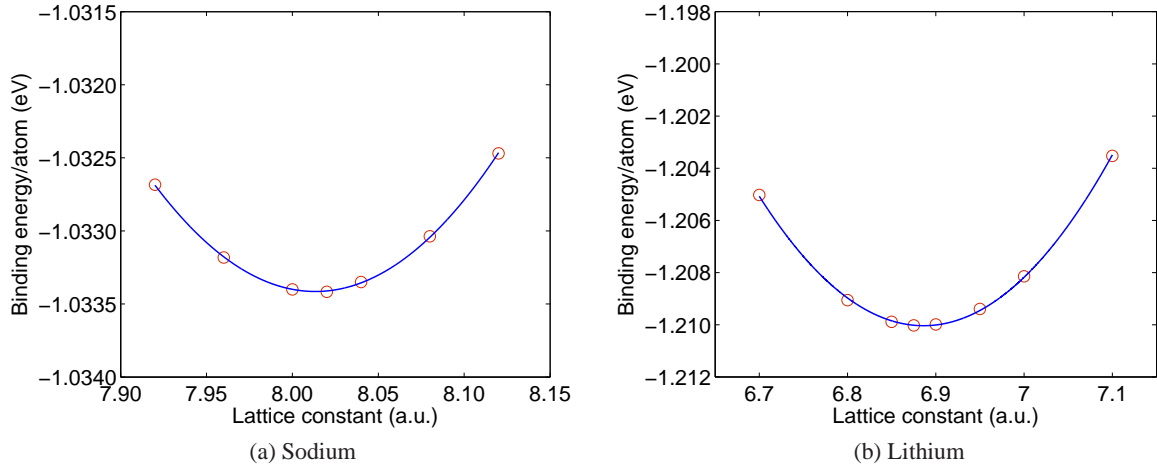


Figure 7: Binding energy as a function of lattice constant

Table 1: Crystalline properties of BCC sodium

Property	LSSGQ	Fiolhais et al. (1995)	Experiment (Fiolhais et al., 1995)
Cohesive energy (eV/atom)	-1.03	-1.02	-1.04
Lattice constant (a. u.)	8.01	8.21	8.21
Bulk modulus (GPa)	5.0	7.1	7.3

periodic solution is attained in the bottommost unit cell of Ω_{RV} . We relax the atoms only along the x_3

Table 2: Crystalline properties of BCC lithium

Property	LSSGQ	Fiolhais et al. (1995)	Experiment (Fiolhais et al., 1995)
Cohesive energy (eV/atom)	-1.21	-1.31	-1.0
Lattice constant (a. u.)	6.87	6.77	6.77
Bulk modulus (GPa)	10.0	14.0	13.3

direction. We evaluate the surface energy using the relation

$$\mathcal{E}_{surface} = \frac{\mathcal{E}_{\Omega_{RV}} - N_e \mathcal{E}_{coh}}{a^2} \quad (73)$$

where $\mathcal{E}_{\Omega_{RV}}$ is the total energy of Ω_{RV} , \mathcal{E}_{coh} is the cohesive energy for a perfect crystal and N_e is the number of electrons in Ω_{RV} .

We begin by verifying the convergence of the calculations with respect to the number of spectral quadrature points (K) in Fig. 8. The rapid convergence, highlighting the efficacy of the method, is noteworthy from the figure. The calculated surface energy and the displacement of the atoms are collected in Table 3. The calculated surface energy is in good agreement with previous values in literature (Vitos et al., 1998). Note that the forces on the atoms in the third layer and beyond are below the threshold value used for the calculations. Also, the convergence of the calculated quantities with respect to size of Ω_{RV} has been verified. Finally, the electron density contours on the mid plane and edge plane are plotted in Fig. 9 for purposes of illustration.

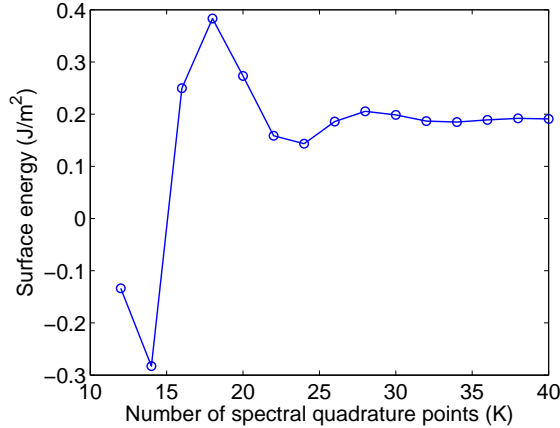


Figure 8: Convergence of LSSGQ method for the (001) surface relaxation of BCC sodium

5. Coarse-Graining Density Functional Theory

Defects in crystalline solids involve elastic and electrostatic fields that may decay extremely slowly. Therefore, accurate computations of such problems require very large computational domains, and even linear-scaling methods can be insufficient to fully understand them. However, the slow decay offers an opportunity to approximate the problem in a manner that reduces the computational complexity with limited loss of accuracy. In this section, we develop a method wherein we solve defects (using DFT) without the

Table 3: Properties of the (001) surface of BCC sodium

Surface energy	0.2 J/m^2
Displacement of atoms	
First layer	0.75 a. u.
Second layer	-0.19 a. u.

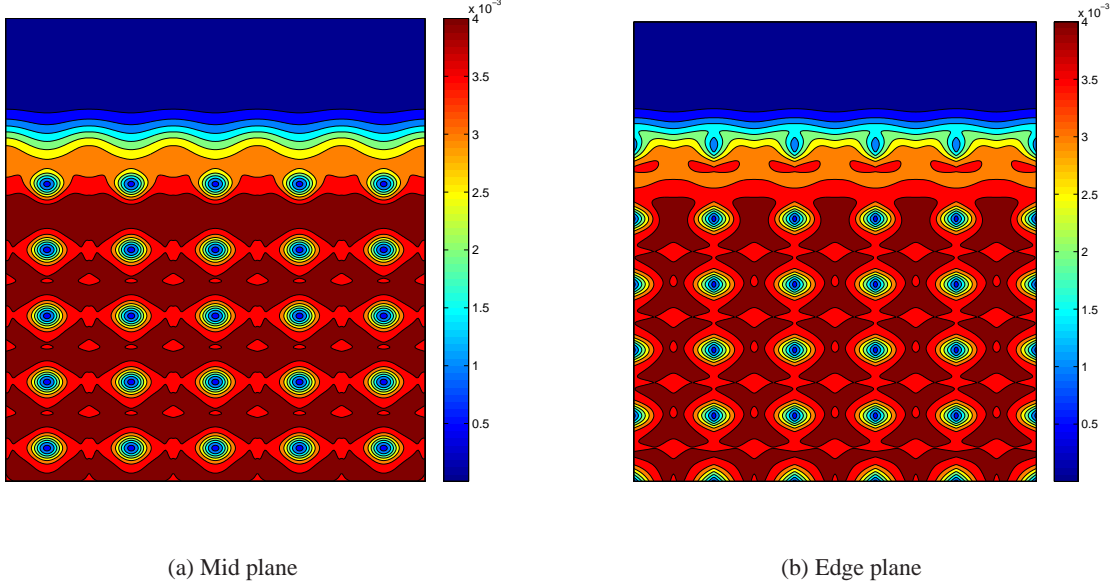


Figure 9: Electron density contours on slices through a (001) surface of BCC sodium

introduction of additional equations or uncontrolled approximations, but with substantially reduced effort. This method is in the spirit of the quasicontinuum approximation used in atomistic models (Tadmor et al. (1996)) and OFDFT (Gavini et al. (2007)), but also has important differences. For definiteness, we discuss the method in terms of the finite-difference approximation scheme. We formulate the coarse-grained DFT (CGDFT) approach in Section 5.1 and validate it through examples in Section 5.2.

5.1. Formulation

Consider the implementation of the LSSGQ method within the context of finite-differences (Section 4.4). Let us denote the collection of all finite-difference nodes $\{\mathbf{x}^p\}_{p=1}^{N_d}$ by \mathcal{N} . In the finite-difference approximation, N_e , U_{band} and S can be represented as the sums of nodal values (cf. Eqns. 61, 63, 64):

$$N_e = h^D \sum_{p=1}^{N_d} \rho^p, \quad U_{band} = \sum_{p=1}^{N_d} u^p, \quad S = \sum_{p=1}^{N_d} s^p, \quad (74)$$

where u and s are used to denote the pointwise band structure energy and pointwise entropy respectively. Further, the nodal values ρ^p , u^p and s^p depend only on values of the spectral quadrature points $\{\lambda_k^p\}_{k=1}^K$ and

weights $\{w_k^p\}_{k=1}^K$ associated with the p^{th} node. (cf. Eqns. 62, 65, 66)

$$\rho^p = \frac{2}{h^D} \sum_{k=1}^K w_k^p g(\lambda_k^p, \lambda_f), \quad (75)$$

$$u^p = 2 \sum_{k=1}^K w_k^p \lambda_k^p g(\lambda_k^p, \lambda_f), \quad (76)$$

$$s^p = -2k_B \sum_{k=1}^K w_k^p [g(\lambda_k^p, \lambda_f) \log g(\lambda_k^p, \lambda_f) + (1 - g(\lambda_k^p, \lambda_f)) \log(1 - g(\lambda_k^p, \lambda_f))]. \quad (77)$$

Finally, recall that we can compute $\{\lambda_k^p\}_{k=1}^K$ and $\{w_k^p\}_{k=1}^K$ associated with a particular node independently of all the other nodes. This observation provides the key idea behind coarse-graining.

When a defect is introduced in an otherwise perfect lattice, we expect complex fluctuations in the electron density and other fields in the vicinity of the defect, but relatively smooth decay away from the defect. Specifically, we expect the atomic displacements to decay in some polynomial fashion and the perturbation in the electron density, pointwise band structure energy and pointwise entropy to decay to zero away from the defect. We wish to take advantage of this long range decay structure to develop an accurate reduced representation. Below, we discuss in detail how this is achieved.

5.1.1. Atomic positions

We select representative atoms and interpolate the displacement for the remaining atoms as is done in the quasicontinuum method (cf., e. g., [Tadmor et al. \(1996\)](#); [Knap and Ortiz \(2001\)](#)). The density of the representative atoms is chosen so as to be high in the vicinity of the defect and to decrease with increasing distance to the defect. This varying resolution can be achieved through a variety of adaptive refinement schemes ([Tadmor et al., 1996](#); [Knap and Ortiz, 2001](#)).

5.1.2. Electron density, band structure energy and entropy

Since we expect the perturbation (due to defect) in the electron density, pointwise band structure energy and pointwise entropy to decay to zero, we seek an approximation scheme with the following decomposition

$$\rho = \rho_0 + \rho_d, \quad u = u_0 + u_d, \quad s = s_0 + s_d \quad (78)$$

where ρ_0, u_0 and s_0 are piecewise periodic, while ρ_d, u_d and s_d decay in a smooth fashion away from the defect. ρ_0, u_0 and s_0 can be determined by an inexpensive periodic calculation. Since ρ_d, u_d and s_d decay smoothly, we can develop an adaptive method for computing them. To this end, we pick representative nodes, the collection of which we denote by \mathcal{R} . The density of finite-difference nodes in \mathcal{R} is chosen to be high in the vicinity of the defect and to progressively decrease with increasing distance from the defect. We then use the LSSGQ method to evaluate the spectral Gauss quadrature points and weights only at the representative nodes $\{\mathbf{x}^p, p \in \mathcal{R}\}$. We then evaluate the Fermi energy by solving for the constraint (Eqn. 74) with the interpolation scheme

$$\rho^p = \rho_0^p + \sum_{q \in \mathcal{R}} \gamma_q^p (\rho^q - \rho_0^q). \quad (79)$$

for the nodes $\{\mathbf{x}^p, p \in \mathcal{N} - \mathcal{R}\}$. γ_q^p are weights decided by the degree of interpolation/approximation used. For the examples in the next section, we use cubic spline interpolations to determine the weights ([Knott \(2000\)](#)). Once we locate the Fermi energy, we calculate the electron density at all the finite-difference nodes

using Eqn. 75 for the nodes $\{\mathbf{x}^p, p \in \mathcal{R}\}$ and Eqn. 79 for the nodes $\{\mathbf{x}^p, p \in \mathcal{N} - \mathcal{R}\}$. The free energy can be determined by evaluating the nodal band structure energy, nodal entropy using Eqns. 76, 77 and using the interpolation

$$u^p = u_0^p + \sum_{q \in \mathcal{R}} \gamma_q^p (u^q - u_0^q) \quad (80)$$

$$s^p = s_0^p + \sum_{q \in \mathcal{R}} \gamma_q^p (s^q - s_0^q). \quad (81)$$

for nodes $\{\mathbf{x}^p, p \in \mathcal{N} - \mathcal{R}\}$. Since we expect ρ_d , u_d and s_d to smoothly decay to zero away from the defect, we can use fewer and fewer representative finite-difference nodes away from the defect to capture these perturbations. As a consequence, the method will display sub-linear scaling with respect to the number of atoms. Therefore, this method opens the possibility of studying systems of much larger size compared to the linear-scaling LSSGQ method.

5.1.3. Electrostatic potential

We write the electrostatic potential

$$\phi = \phi_0 + \phi_d \quad (82)$$

and evaluate the perturbation ϕ_d by solving the equation

$$-\frac{1}{4\pi} \nabla^2 \phi_d = \rho - \rho_0 + b - b_0. \quad (83)$$

By its very nature, ϕ_d is localized near the defect and decays to zero away from it. Therefore, it can be easily calculated by using any adaptive discretization tailored to this feature.

We summarize the CGDFT method in Algorithm 2.

5.2. Validation

In this section, we verify the CGDFT method by applying it to a one-dimensional model problem and a Kohn-Sham surface relaxation problem, in Sections 5.2.1 and 5.2.2 respectively.

5.2.1. One-dimensional model

We study the efficiency of CGDFT in the framework of the one-dimensional model presented in Section 4.5.1. Consider a chain consisting of 101 atoms with unit spacing. We introduce a point defect by removing the center atom. We coarse-grain in both directions from the position of the defect and adopt the procedure described in Section 5.1 to solve for the electron density and the energy. We plot the error in the defect energy and defect electron density so obtained, as a function of the number of representative nodes in Fig. 10. The error is measured and normalized with respect to the fully resolved LSSGQ solution. It is clear that we get rapid convergence in the solution for both metals and insulators. As expected, convergence in the case of insulators is much more rapid owing to the highly localized nature of the defect perturbation. Therefore, CGDFT can be utilized to solve the defect problem at a small fraction of the original computational cost, without any significant loss of accuracy.

Algorithm 2: CGDFT method

Generate guess for positions of the representative atoms (\mathbf{R})

repeat Relaxation of representative atoms

 Calculate charge density of the nuclei (b)

 Generate piecewise periodic ρ_0 and ϕ_0

 Generate guess for electron density (ρ)

repeat Self-Consistent loop: SCF

 Calculate electrostatic potential (ϕ) by solving Eqn. 83

 Calculate exchange correlation potential (V_{xc})

 Calculate spectral quadrature points and weights at representative finite-difference nodes

 Calculate Fermi energy (λ_f) using Eqn. 74 (utilizing interpolation given by Eqn. 79)

 Calculate electron density at representative nodes ($\rho^p, p \in \mathcal{R}$) using Eqn. 75

 Interpolate the electron density to nodes ($\rho^p, p \in \mathcal{N} - \mathcal{R}$) using Eqn. 79

 Update the electron density (mixing)

until Convergence of self-consistent iteration;

 Calculate the forces on the nuclei using Eqn. 21

until Energy minimized with respect to positions of representative atoms;

Evaluate free energy \mathcal{F} using Eqns. 74, 75, 76, 77, 80 and 81 for the band structure energy and entropy.

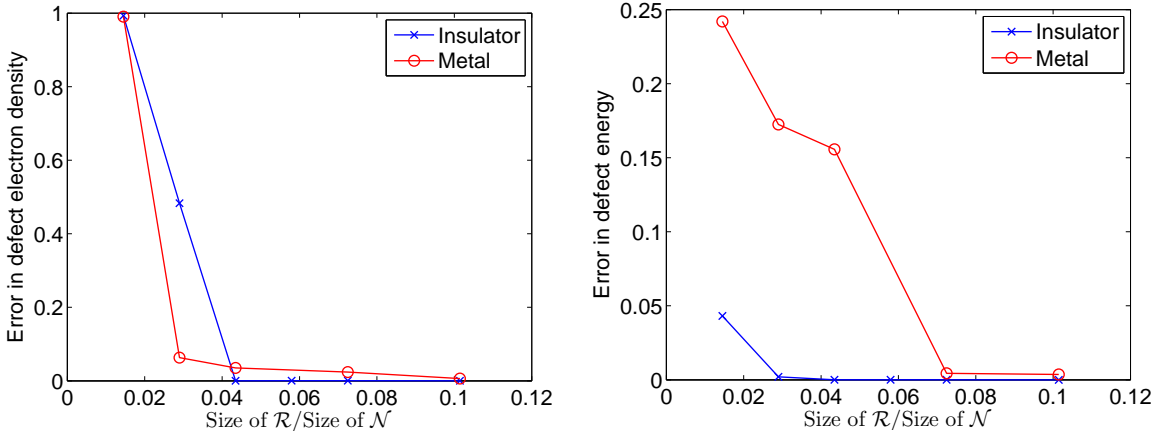


Figure 10: Convergence of the CGDFT method for the one-dimensional model problem with a point defect

5.2.2. Kohn-Sham problem: Surface relaxation

We now study the (001) surface relaxation of BCC sodium using CGDFT, previously analyzed using the LSSGQ method in Section 4.5.2. We employ the procedure described in Section 4.5.2 with the introduction of coarse-graining described in Section 5.1. However, we do not introduce coarse-graining for the electrostatic potential since the computational time involved in its calculation is negligible. We compare the results obtained using CGDFT with the fully resolved solution obtained using the LSSGQ method. A summary of the results is presented in Table 4. We can see that there are significant computational savings when using CGDFT at no appreciable loss of accuracy. It should be carefully noted that the surface relaxation problem is effectively one-dimensional for purposes of coarse-graining. As a consequence, the computational savings

accrued are likely to be less substantial than in fully two or three dimensional problems.

Table 4: Summary of the results for CGDFT

Number of representative nodes/Total number of nodes	0.25
Number of representative atoms/Total number of atoms	0.3
Normalized error in surface energy	0.12 %
Normalized error in displacement of atoms	
First layer	0.11 %
Second layer	0.17 %

6. Conclusions

We have developed a real-space formulation for coarse-graining DFT (CGDFT). Traditional implementations of DFT solve for the orbitals, which results in cubic-scaling with respect to the number of atoms. Furthermore, they are not amenable to coarse-graining because of the global nature of the orthonormality constraint. In order to overcome this difficulty, we have developed a linear-scaling method for DFT (LSSGQ), where we directly evaluate the electron density without evaluating the individual orbitals. We accomplish this direct evaluation by performing Gauss quadrature over the spectrum of the linear Hamiltonian operator in each iteration of the SCF method. As a second approximation step, we have appended coarse-graining approximations to the LSSGQ method. Specifically, we decompose the solution into a periodic component and the perturbation caused by lattice defects. Since the perturbation is expected to decay rapidly away from the defect, fine resolution is only needed in the vicinity of the defect and the resolution can be coarse-grained elsewhere. This coarse-graining enables the analysis of defects at a fraction of the computational cost, without any appreciable loss of accuracy. We have validated the LSSGQ and CGDFT methods through a number of examples of application, the results of which are in good agreement with literature values. Furthermore, we have verified the convergence of the CGDFT solution to the fully resolved solution. These verification and validation examples, taken together, show CGDFT to be highly transferable, efficient and accurate.

We close by pointing to some outstanding issues and limitations of the CGDFT method and to possible extensions of the method thereof. A principal shortcoming of the implementation described in this paper is the non-automated nature of the coarse-graining approximations. An optimal and automated coarse-graining technique for CGDFT would greatly extend the range, scope and robustness of the method. In addition, a Rayleigh-Ritz or Galerkin spatial discretization scheme, e. g., based on finite elements, would endow the approximation scheme with a variational structure, thereby opening the way to energy-based adaptive methods. In addition, the convergence of variational approximation schemes can be established by means of powerful tools such as Γ -convergence. However, it should be noted that finite-element bases lack orthogonality in general and result in generalized eigenvalue problems. This property may potentially add to the computational cost and implementational complexity associated with the method. These issues are currently being investigated by the authors.

Acknowledgements

This work draws from the thesis of PS at the California Institute of Technology. We gratefully acknowledge the support of the US Army Research Office (under MURI grant number W911NF-07-1-0410),

the US Department of Energy National Nuclear Security Administration (under Award Number DE-FC52-08NA28613 through Caltech's ASC/PSAAP Center for the Predictive Modeling and Simulation of High Energy Density Dynamic Response of Materials) and the US National Science Foundation (under PIRE grant number OISE-0967140).

A. Padé approximation and recursion method

The recursion method (Haydock, 1980) has found many applications in the physics literature. Here, we discuss the close connection between the LSSGQ and Recursion methods. In doing so, we also establish their relationship to the Padé approximation. Details of the notation not described here can be found in Sections 3 and 4.

Let $\mathcal{R}_{\mathcal{H}}(z) = (z\mathcal{I} - \mathcal{H})^{-1}$ represent the resolvent of the operator \mathcal{H} for $z \in \rho(\mathcal{H}) = \mathbb{C} \setminus \sigma(\mathcal{H})$ where $\rho(\mathcal{H})$ denotes the resolvent set of \mathcal{H} and \mathbb{C} is the set of all complex numbers. $\mathcal{R}_{\mathcal{H}}(z)$ is holomorphic $\forall z \in \rho(\mathcal{H})$. For $z \in \rho(\mathcal{H})$, $\zeta \in H$ it follows

$$\mathcal{R}_{\mathcal{H}}(z) = \int_{\sigma(\mathcal{H})} \frac{1}{z - \lambda} d\mathcal{E}(\lambda), \quad (84)$$

$$G_{\zeta, \zeta}(z) = (\mathcal{R}_{\mathcal{H}}(z)\zeta, \zeta) = \int_{\sigma(\mathcal{H})} \frac{1}{z - \lambda} d\mathcal{E}_{\zeta, \zeta}(\lambda). \quad (85)$$

Taking the series expansion of $G_{\zeta, \zeta}(z)$ around the point $z = \infty$, we obtain

$$\begin{aligned} G_{\zeta, \zeta}(z) &= \sum_{k=0}^{\infty} \frac{1}{z^{k+1}} (\mathcal{H}^k \zeta, \zeta) \\ &= \sum_{k=0}^{\infty} \frac{1}{z^{k+1}} \int_a^b \lambda^k d\mathcal{E}_{\zeta, \zeta}(\lambda) = \sum_{k=0}^{\infty} \frac{\mu_k}{z^{k+1}} \end{aligned} \quad (86)$$

which converges for $|z| > \|\mathcal{H}\|$.

Diagonal Padé approximants for $G_{\zeta, \zeta}(z)$ are rational functions of the form $q_K(z)/p_K(z)$ which satisfy the following property

$$\left(G_{\zeta, \zeta}(z) - \frac{q_K(z)}{p_K(z)} \right) = \mathcal{O} \left(\frac{1}{z^{2K+1}} \right) \quad (87)$$

and are locally the best rational approximations for a given power series like Eqn. 86 (cf., e. g., Suetin (2002)). The denominator polynomials $p_K(z)$, are a set of orthogonal polynomials with respect to the measure $\mathcal{E}_{\zeta, \zeta}$ (cf., e. g., Moren and Branquinho (2008))

$$\int_a^b \lambda^k p_K(\lambda) d\mathcal{E}_{\zeta, \zeta}(\lambda) = 0, \quad k = 0, 1, \dots, K-1 \quad (88)$$

and the numerator polynomial $q_K(z)$ can be expressed in terms of $p_K(z)$ as follows

$$q_K(z) = \int_a^b \frac{p_K(z) - p_K(\lambda)}{z - \lambda} d\mathcal{E}_{\zeta, \zeta}(\lambda). \quad (89)$$

The Padé approximants are constructed using the coefficients of the power series and provide an efficient analytic continuation of the series beyond the domain of convergence, which in the present case is given

by $|z| > \|\mathcal{H}\|$. The convergence of diagonal Padé approximants follows from Markov's theorem (cf., e. g., [Suetin \(2002\)](#)), whereby any Markov function like $G_{\zeta,\zeta}(z)$ can be recovered in $\rho(\mathcal{H})$ from the coefficients of the Laurent expansion of the function at the point $z = \infty$ (i.e. from the moments of the measure $\mathcal{E}_{\zeta,\zeta}$).

The orthogonal polynomials $\{p_k\}_{k=1}^K$ can be generated via the recurrence relation (cf., e. g., [Golub and Meurant \(2010\)](#))

$$\begin{aligned} p_{k+1}(\lambda) &= (\lambda - a_{k+1})p_k(\lambda) - b_k^2 p_{k-1}(\lambda), \quad k = 0, 1, \dots, K-1 \\ p_{-1}(\lambda) &= 0, \quad p_0(\lambda) = 1 \end{aligned} \quad (90)$$

where

$$\begin{aligned} a_{k+1} &= \frac{\langle \lambda p_k, p_k \rangle_{\zeta}}{\langle p_k, p_k \rangle_{\zeta}}, \quad k = 0, 1, \dots, K-1 \\ b_k^2 &= \frac{\langle p_k, p_k \rangle_{\zeta}}{\langle p_{k-1}, p_{k-1} \rangle_{\zeta}}, \quad k = 1, 2, \dots, K-1. \end{aligned} \quad (91)$$

Using Eqns. 88 and 89, we see that $\{q_k\}_{k=1}^K$ can be generated with a similar recurrence relation as that for $\{p_k\}_{k=1}^K$ but with different initializing conditions

$$\begin{aligned} q_{k+1}(\lambda) &= (\lambda - a_{k+1})q_k(\lambda) - b_k^2 q_{k-1}(\lambda), \quad k = 0, 1, \dots, K-1 \\ q_{-1}(\lambda) &= -1, \quad q_0(\lambda) = 0 \text{ and } b_0^2 = 1. \end{aligned} \quad (92)$$

Corresponding to these recurrence relations we have the tridiagonal matrix

$$J_K = \begin{pmatrix} a_1 & 1 & & & \\ b_1^2 & a_2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & b_{K-2}^2 & a_{K-1} & 1 \\ & & & b_{K-1}^2 & a_K \end{pmatrix} \quad (93)$$

Expanding $\det(zI_k - J_k)$ with respect to its last row, we obtain the recurrence formula for the determinant

$$\begin{aligned} \det(zI_{k+1} - J_{k+1}) &= (z - a_{k+1})\det(zI_k - J_k) - b_k^2 \det(zI_{k-1} - J_{k-1}), \quad k = 0, 1, \dots, K-1 \\ \det(zI_{-1} - J_{-1}) &= 0, \quad \det(zI_0 - J_0) = 1 \end{aligned} \quad (94)$$

which is the same as that satisfied by $\{p_k\}_{k=0}^{K-1}$ as well as $\{\det(zI_k - \hat{J}_k)\}_{k=0}^{K-1}$. Consequently

$$p_K(z) = \det(zI_K - J_K) = \det(zI_K - \hat{J}_K). \quad (95)$$

Similarly

$$q_K(z) = \det(zI_{K-1} - J_{K-1}^1) = \det(zI_{K-1} - \hat{J}_{K-1}^1) \quad (96)$$

where the superscript i is used to denote the matrix obtained by deleting the first i rows and columns. Therefore

$$\frac{q_K(z)}{p_K(z)} = \frac{\det(zI_{K-1} - \hat{J}_{K-1}^1)}{\det(zI_K - \hat{J}_K)} = e_1(zI_K - \hat{J}_K)^{-1}e_1. \quad (97)$$

Above, the first column of I_K is denoted by e_1 .

Expanding $\det(zI_K - \hat{J}_K)$ about the first row or column we obtain from Eqn. 97

$$\frac{q_K(z)}{p_K(z)} = \frac{1}{z - a_1 - \frac{\det(zI_{K-2} - \hat{J}_{K-2}^2)}{\det(zI_{K-2} - \hat{J}_{K-2}^1)}} \quad (98)$$

and continuing similarly we obtained a continued fraction representation for the Padé approximant

$$\frac{q_K(z)}{p_K(z)} = \frac{1}{z - a_1 - \frac{b_1^2}{z - a_2 - \dots - \frac{b_{K-1}^2}{z - a_K}}} \quad (99)$$

In the recursion method, the above expression is used to find the projected density of states (Haydock, 1980) from which all the necessary quantities are evaluated. Since the projected density of states has poles at the zeros of $p_K(z)$, a number of techniques to smoothen it have been developed. These include either using a terminator, simplest one being $\{a_k\}_{k=K+1}^\infty = a_\infty$, $\{b_k\}_{k=K}^\infty = b_\infty$ or evaluating the Padé approximant slightly away from the real axis. However, these are uncontrolled approximations and can sometimes lead to inaccuracies.

Finally, we look at the connection between Padé approximation and Gauss quadrature (cf., e. g., Van Assche (2006)). Multiplying both sides of Eqn. 87 by a polynomial of degree at most $2K - 1$ denoted by $\pi_{2K-1}(z)$, integrating along a contour Γ encircling the real line and subsequently using Eqn. 85 we obtain

$$\int_a^b \pi_{2K-1}(\lambda) d\mathcal{E}_{\eta,\eta}(\lambda) = \sum_{j=1}^K w_j \pi_{2K-1}(\lambda_j) \quad (100)$$

where

$$w_j = \frac{q_K(\lambda_j)}{p'_K(\lambda_j)} \quad (101)$$

is the residue of the Padé approximant at the zeros λ_j of p_K . It is clear that all polynomials of degree $2K - 1$ are integrated exactly, therefore establishing the connection between the Padé approximation, recursion method and the LSSGQ method.

B. LSSGQ method applied to systems with periodicity

In this section, we discuss the application of the LSSGQ method to the study of periodic systems. Typically, this would involve the need for Bloch periodic boundary conditions on the orbital. However, since we directly evaluate the electron density and not the individual orbitals, we can circumvent the need for Bloch periodic boundary conditions using the procedure described below. For clarity, we discuss the procedure in terms of the finite-difference approximation. However, the method is not restricted by the choice of basis.

First, we specify a representative volume Ω_{RV} which has the property that the solution for the entire system can be ascertained through a translational mapping of the solution obtained in Ω_{RV} . We use a finite-difference grid of uniform spacing h to discretize Ω_{RV} , which we denote by χ_{RV} . Next, we define an extended volume $\Omega_{EV} \supset \Omega_{RV}$, which is discretized using a finite-difference grid of spacing h , denoted by $\chi_{EV} \supset \chi_{RV}$. We evaluate the spectral quadrature nodes and weights for χ_{RV} using the procedure outlined

in Section 4.4. It is important to note that the size of Ω_{EV} is chosen such that all the vectors in the Lanczos iteration given by Eqn. 59 have nonzero components only for finite-difference nodes inside Ω_{EV} . Since the initial vector \mathbf{v}_0 always has a single nonzero entry, the required size of Ω_{EV} can be easily calculated using the bandwidth of the discretized \mathcal{H} and the number of spectral quadrature points used.

References

- Bachelet, G. B., Hamann, D. R., Schlüter, M., 1982. Pseudopotentials that work: From H to Pu. *Physical Review B* 26 (8), 4199–4228.
- Barrault, M., Cancès, E., Hager, W., Bris, C. L., 2007. Multilevel domain decomposition for electronic structure calculations. *Journal of Computational Physics* 222 (1), 86 – 109.
- Bernstein, N., Kermode, J. R., Csanyi, G., 2009. Hybrid atomistic simulation methods for materials systems. *Reports on Progress in Physics* 72 (2), 026501 (25pp).
- Bowler, D. R., Miyazaki, T., 2012. $\mathcal{O}(N)$ methods in electronic structure calculations. *Reports on Progress in Physics* 75 (3), 036503.
- Cancès, E., Bris, C. L., Lions, P.-L., 2008. Molecular simulation and related topics: some open mathematical problems. *Nonlinearity* 21 (9), T165.
- Castro, A., Appel, H., Oliveira, M., Rozzi, C. A., Andrade, X., Lorenzen, F., Marques, M. A. L., Gross, E. K. U., Rubio, A., 2006. octopus: a tool for the application of time-dependent density functional theory. *Physica Status Solidi B-Basic Solid State Physics* 243 (11), 2465–2488.
- Chelikowsky, J. R., Troullier, N., Saad, Y., 1994. Finite-difference-pseudopotential method: Electronic structure calculations without a basis. *Physical Review Letters* 72 (8), 1240–1243.
- Choly, N., Lu, G., E, W., Kaxiras, E., Mar 2005. Multiscale simulations in simple metals: A density-functional-based methodology. *Phys. Rev. B* 71 (9), 094101.
- Cuitino, A., Stainier, L., Wang, G., Strachan, A., Cagin, T., Goddard, W., Ortiz, M., Jan 2002. A multiscale approach for modeling crystalline solids. *J Comput-Aided Mater* 8 (2-3), 127–149.
- Fang, H.-r., Saad, Y., MAR 2009. Two classes of multisection methods for nonlinear acceleration. *NUMERICAL LINEAR ALGEBRA WITH APPLICATIONS* 16 (3), 197–221.
- Finnis, M., 2003. Interatomic forces in condensed matter. Oxford University Press.
- Fiolhais, C., Perdew, J. P., Armster, S. Q., MacLaren, J. M., Brajczewska, M., 1995. Dominant density parameters and local pseudopotentials for simple metals. *Physical Review B* 51 (20), 14001–14011.
- Forsythe, G., Malcolm, M., Moler, C., 1973. Computer methods for mathematical computations. Prentice-Hall.
- Galli, G., Parrinello, M., Dec 1992. Large scale electronic structure calculations. *Phys. Rev. Lett.* 69 (24), 3547–3550.

- Garcia-Cervera, C. J., Lu, J., Xuan, Y., E, W., 2009. Linear-scaling subspace-iteration algorithm with optimally localized nonorthogonal wave functions for kohn-sham density functional theory. *Physical Review B (Condensed Matter and Materials Physics)* 79 (11), 115110.
- García-Cervera, C. J., Lu, J., Xuan, Y., E, W., Mar 2009. Linear-scaling subspace-iteration algorithm with optimally localized nonorthogonal wave functions for kohn-sham density functional theory. *Phys. Rev. B* 79 (11), 115110.
- Gavini, V., Bhattacharya, K., Ortiz, M., Jan 2007. Quasi-continuum orbital-free density-functional theory: A route to multi-million atom non-periodic dft calculation. *J Mech Phys Solids* 55 (4), 697–718.
- Gillan, M. J., 1989. Calculation of the vacancy formation energy in aluminium. *Journal of Physics: Condensed Matter* 1 (4), 689.
- Goedecker, S., Jul 1999. Linear scaling electronic structure methods. *Rev. Mod. Phys.* 71 (4), 1085–1123.
- Golub, G., Meurant, G., 2010. *Matrices, Moments and Quadrature with Applications*. Princeton University Press.
- Gonze, X., Beuken, J. M., Caracas, R., Detraux, F., Fuchs, M., Rignanese, G. M., Sindic, L., Verstraete, M., Zerah, G., Jollet, F., Torrent, M., Roy, A., Mikami, M., Ghosez, P., Raty, J. Y., Allan, D. C., 2002. First-principles computation of material properties: the abinit software project. *Computational Materials Science* 25, 478–492(15).
- Govind, N., Wang, Y. A., Carter, E. A., 1999. Electronic-structure calculations by first-principles density-based embedding of explicitly correlated systems. *The Journal of Chemical Physics* 110 (16), 7677–7688.
- Haydock, R., 1980. *Solid State Physics*. Vol. 35. Academic Press.
- Haydock, R., Heine, V., Kelly, M., Jan 1972. Electronic-structure based on local atomic environment for tight-binding bands. *J Phys C Solid State* 5 (20), 2845–&.
- Haydock, R., Heine, V., Kelly, M., Jan 1975. Electronic-structure based on local atomic environment for tight-binding bands: Ii. *J Phys C Solid State* 8 (16), 2591–2605.
- Hohenberg, P., Kohn, W., 1964. Inhomogeneous electron gas. *Physical Review* 136 (3B), B864–B871.
- Ismail-Beigi, S., Arias, T. A., 2000. New algebraic formulation of density functional calculation. *Computer Physics Communications* 128 (1-2), 1 – 45.
- Kleinman, L., Bylander, D. M., 1982. Efficacious form for model pseudopotentials. *Physical Review Letters* 48 (20), 1425–1428.
- Knap, J., Ortiz, M., SEP 2001. An analysis of the quasicontinuum method. *JOURNAL OF THE MECHANICS AND PHYSICS OF SOLIDS* 49 (9), 1899–1923.
- Knott, G. D., 2000. *Interpolating Cubic Splines*. Springer-Verlag.
- Kohn, W., Sham, L. J., 1965. Self-consistent equations including exchange and correlation effects. *Physical Review* 140 (4A), A1133–A1138.

- Kresse, G., Furthmüller, J., 1996. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical Review B* 54 (16), 11169–11186.
- Langreth, D. C., Mehl, M. J., 1983. Beyond the local-density approximation in calculations of ground-state electronic properties. *Physical Review B* 28 (4), 1809–1834.
- Lu, G., Tadmor, E. B., Kaxiras, E., 2006. From electrons to finite elements: A concurrent multiscale approach for metals. *Physical Review B (Condensed Matter and Materials Physics)* 73 (2), 024108.
- Martin, R., 2004. *Electronic Structure: Basic theory and practical methods*. Cambridge University Press.
- Mauri, F., Galli, G., Car, R., Apr 1993. Orbital formulation for electronic-structure calculations with linear system-size scaling. *Phys. Rev. B* 47 (15), 9973–9976.
- Miller, R. E., Tadmor, E. B., Jan 2009. A unified framework and performance benchmark of fourteen multi-scale atomistic/continuum coupling methods. *Model Simul Mater Sc* 17 (5), 053001.
- Moren, A., Branquinho, A., 2008. *Coimbra lecture notes on Orthogonal Polynomials*. Nova Science Publishers.
- Parr, R., Yang, W., 1989. *Density-functional theory of atoms and molecules*. Oxford University Press.
- Pask, J. E., Klein, B. M., Fong, C. Y., Sterne, P. A., 1999. Real-space local polynomial basis for solid-state electronic-structure calculations: A finite-element approach. *Physical Review B* 59 (19), 12352–12358.
- Perdew, J. P., Chevary, J. A., Vosko, S. H., Jackson, K. A., Pederson, M. R., Singh, D. J., Fiolhais, C., 1992. Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation. *Physical Review B* 46 (11), 6671–6687.
- Phillips, R., 2001. *Crystals, defects and microstructure: modeling across scales*. Cambridge University Press.
- Rappe, A. M., Rabe, K. M., Kaxiras, E., Joannopoulos, J. D., 1990. Optimized pseudopotentials. *Physical Review B* 41 (2), 1227–1230.
- Rudin, W., 1991. *Functional Analysis*. Mc Graw Hill.
- Saad, Y., Schultz, M. H., 1986. GMRES: A Generalized Minimal Residual Algorithm for Solving Nonsymmetric Linear Systems. *SIAM Journal on Scientific and Statistical Computing* 7 (3), 856–869.
- Segall, M. D., Lindan, P. J. D., Probert, M. J., Pickard, C. J., Hasnip, P. J., Clark, S. J., Payne, M. C., 2002. First-principles simulation: ideas, illustrations and the CASTEP code. *Journal of Physics: Condensed Matter* 14 (11), 2717–2744.
- Skylaris, C.-K., Haynes, P. D., Mostofi, A. A., Payne, M. C., 2005. Introducing ONETEP: Linear-scaling density functional simulations on parallel computers. *The Journal of Chemical Physics* 122 (8), 084119.
- Suetin, S. P., 2002. Padé approximants and efficient analytic continuation of a power series. *Russian Mathematical Surveys* 57 (1), 43.
- Suryanarayana, P., Bhattacharya, K., Ortiz, M., 2011. A mesh-free convex approximation scheme for kohn-sham density functional theory. *Journal of Computational Physics* 230 (13), 5226 – 5238.

- Suryanarayana, P., Gavini, V., Blesgen, T., Bhattacharya, K., Ortiz, M., 2010. Non-periodic finite-element formulation of kohn-sham density functional theory. *Journal of the Mechanics and Physics of Solids* 58 (2), 256 – 280.
- Tadmor, E., Ortiz, M., Phillips, R., JUN 1996. Quasicontinuum analysis of defects in solids. *PHILOSOPHICAL MAGAZINE A-PHYSICS OF CONDENSED MATTER STRUCTURE DEFECTS AND MECHANICAL PROPERTIES* 73 (6), 1529–1563.
- Troullier, N., Martins, J. L., 1991. Efficient pseudopotentials for plane-wave calculations. *Physical Review B* 43 (3), 1993–2006.
- Tsuchida, E., 2004. Ab initio molecular-dynamics study of liquid formamide. *The Journal of Chemical Physics* 121 (10), 4740–4746.
- Van Assche, W., 2006. Padé and hermite-padé approximation and orthogonality. *Surveys in Approximation Theory* (2006) 2, 61–91.
- Vanderbilt, D., 1990. Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. *Physical Review B* 41 (11), 7892–7895.
- Vitos, L., Ruban, A., Skriver, H., KollÅar, J., 1998. The surface energy of metals. *Surface Science* 411 (1-2), 186 – 202.
- Vogel, C. R., 2002. Computational methods for inverse problems. *Society for Industrial and Applied Mathematics*.
- Weinert, M., Davenport, J. W., Jun 1992. Fractional occupations and density-functional energies and forces. *Phys. Rev. B* 45 (23), 13709–13712.